

# Sampling and Reconstruction of Sparse Signals on Circulant Graphs - An Introduction to Graph-FRI

M. S. Kotzagiannidis\*, P. L. Dragotti

*Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK*

---

## Abstract

With the objective of employing graphs toward a more generalized theory of signal processing, we present a novel sampling framework for (wavelet-)sparse signals defined on circulant graphs which extends basic properties of Finite Rate of Innovation (FRI) theory to the graph domain, and can be applied to arbitrary graphs via suitable approximation schemes. At its core, the introduced Graph-FRI-framework states that any  $K$ -sparse signal on the vertices of a circulant graph can be perfectly reconstructed from its dimensionality-reduced representation in the Graph Fourier domain of minimum size  $2K$ . By leveraging the recently developed theory of e-splines and e-spline wavelets on graphs, one can decompose this graph spectral transformation into the multiresolution low-pass filtering operation with a graph e-spline filter, and subsequent transformation to the spectral graph domain; this allows to infer a distinct sampling pattern, and, ultimately, the structure of an associated coarsened graph, which preserves essential properties of the original, including circularity and, where applicable, the graph generating set.

*Keywords:* graph signal processing, sampling on graphs, sparse sampling, graph wavelet, finite rate of innovation

---

## 1. Introduction

Contributions to graph signal processing (GSP) theory have aspired to create extensions of traditional signal processing notions to the graph domain, motivated by the need to gain a deeper understanding of how the complex connectivity of graphs may be leveraged for more sophisticated processing, computational efficiency and superior performance, all the while heading toward a more generalized theory of SP [1]. The inherent challenge of interpreting and incorporating newly arising data dependencies, while maintaining equivalencies to classical cases, has given rise to a variety of different approaches, borrowing notions from i.a. algebraic and spectral graph theory [2], algebraic signal processing theory [3], and general matrix theory [4].

A breadth of intriguing GSP problems such as graph wavelet analysis ([5], [6], [7], [8], [9]), graph signal interpolation and recovery ([10], [11], [12]), up to graph-based image

---

\*Corresponding author

*Email address:* m.kotzagiannidis13@ic.ac.uk (M. S. Kotzagiannidis)

processing ([13], [7]) and semi-supervised learning ([14], [15]), have been derived in the wake of two elementary model assumptions for the central graph operator: the positive semi-definite *graph Laplacian matrix*, and the more generalized *graph adjacency matrix*. In search of concrete analogies between traditional and graph SP, the class of circulant graphs has been noted for its linear shift invariance property and provided the foundation for intuitive graph signal sampling and filtering operations, as first established in ([16], [6], [17]), not least of all due to its characterization by the (permuted) DFT in the spectral graph domain. This previously inspired our derivation of families of signal-sparsifying, critically-sampled and vertex-localized graph spline and graph e-spline wavelet filterbanks on circulant graphs ([18], [19], [20]), with the vanishing moment property of the graph Laplacian operator and its parameterised generalization, the e-graph Laplacian, at its core. Equipped with reproduction and annihilation properties for higher-order complex exponential polynomial graph signals, these filterbanks can be applied iteratively for a sparse multiresolution signal representation on suitable coarsened graphs.

In an effort to further pursue a widespread motivation to elucidate sparsity on graphs, this work addresses the problem of sparse sampling and coarsening on graphs by proposing an intuitive framework for circulant graphs, as an extension to classical approaches in the Euclidean domain, which can be generalized to arbitrary graphs by using the former as building blocks.

Complementary to our discussion of (e-)spline wavelets on circulant graphs in [20], we proceed to investigate the sampling and recovery of sparse, and hence, wavelet-sparse graph signals on circulant graphs, for which we present a novel framework as an extension of the traditional Finite Rate of Innovation (FRI) theory ([21], [22]) to the graph domain. Given the dimensionality-reduced, graph Laplacian-based spectral representation  $\mathbf{y}$  of a sparse graph signal  $\mathbf{x}$  defined on the vertices of a circulant graph, we show that the latter can be perfectly recovered using Prony’s method [22], while identifying the coarse graph associated with the vertex-localized version of  $\mathbf{y}$  through i.a. a scheme of spectral sampling; we additionally refine and extend this approach to encompass ‘compressible’ graph signals, which have a sparse graph wavelet representation in light of our newly derived constructions, including the sets of (piecewise) polynomials and complex exponential polynomials. Eventually, generalizations to (multi-dimensional) sampling on arbitrary graphs can be made i.a. on the basis of graph products, as we previously demonstrated for graph spline wavelet analysis [20].

### Related Work:

Signal recovery on graphs, denoting more broadly the empirical study as opposed to the analytical framework, has been tackled i.a. under the premise that a signal is smooth with respect to the underlying graph, and can for instance be formulated as an optimization problem in different settings ([12], [23]). In [24], [25], [26], and [27], sampling theory for graphs, providing the specialized and more rigorous theorization of the former, is explored with predominant regard to the subspace of bandlimited graph signals under different assumptions; hereby, [26] and [27] provide two alternative interpretations of bandlimitedness in the graph domain, where, in particular, the latter uses matrix algebra to establish a linear reconstruction approach, based on the knowledge and suitable choice of the retained sample locations. Moving beyond the traditional domain, sampling theory in the context of graphs has furthermore attempted to address graph coarsening, as can be seen in [27], also a problem in itself ([28], [29]), which bears the challenge of

identifying a meaningful underlying graph for the sampled signal and has been featured to a lesser extent.

In particular, the graph coarsening scheme introduced in [27] is comparable to the coarsening approach via spectral sampling in our framework, up to the choice of the sampling set and resulting property preservations, and with the further distinction that we iteratively filter the given graph signal with a suitable graph e-spline filter. While the former requires  $K$  entries of suitably chosen sample locations (for bandlimited signals with bandwidth  $K$ ) for perfect recovery, our downsampling pattern is fixed and primarily used to extract the coarsened graph corresponding to the sampled graph signal, under preservation of certain graph properties, as well as independent of the reconstruction scheme, which solely requires the input of the dimensionality-reduced spectral graph signal. In addition, we consider sparse and graph wavelet-sparse, as opposed to bandlimited, graph signals, encompassing a wide variety of graph signal classes, which do not necessarily belong to a fixed subspace, as facilitated through suitable graph spline wavelet analysis. We first investigated the problem of sparse signal reconstruction on circulant graphs in [30] in the context of noisy recovery under (graph-)perturbations, with a preliminary discussion in [19].

Due to our focus on sparse graph signals, the comparison with compressive sensing (CS) [31] is imperative. In CS theory, a sparse signal  $\mathbf{x} \in \mathbb{R}^N$  can be recovered with high probability from the dimensionality-reduced, sampled signal  $\mathbf{y} = \mathbf{A}\mathbf{x}$  under suitable conditions on the rectangular sampling operator  $\mathbf{A} \in \mathbb{R}^{M \times N}$  with  $M \ll N$  and sparsity  $K = \|\mathbf{x}\|_0$ , by solving an  $l_1$ -minimization problem, or alternatively, using greedy reconstruction algorithms [32]. While in contrast to compressive sensing approaches [33], the recovery of the sparse vector  $\mathbf{x}$  in our scheme is exact at the critical dimension of  $2K$  measurements and based on a direct, spectral estimation technique, known as Prony's method ([34], [22]), we note that neither requires knowledge of the locations of the non-zero entries. In addition, compressed sensing theory can be extended to the recovery of non-sparse signals  $\mathbf{x} = \mathbf{D}\mathbf{c}$  that have a sparse representation  $\mathbf{c}$  in properly designed, over-complete dictionaries  $\mathbf{D}$  [35], which has also been addressed in the context of graphs by training a graph-based dictionary [23]. Our sampling framework takes a similar approach in that smooth (wavelet-sparse) graph signals  $\mathbf{x}$  are filtered with a circulant multilevel graph wavelet transform in order to produce sparse signals  $\mathbf{c}$  which can subsequently be sampled; nevertheless, the recovery of  $\mathbf{x}$  from  $\mathbf{c}$  ultimately follows from the invertibility of the wavelet transform.

In this work, we make the following main *contributions*:

- A novel framework for the sampling and perfect reconstruction of sparse and graph-wavelet-sparse signals on circulant graphs (Thm. 4.1)
- A general scheme to extract the coarse graph associated with the sampled signal to accompany the above, including a property-preserving approach based on spectral sampling (Thm 4.2)
- Generalizations to sampling and recovery on arbitrary graphs via graph product decomposition approximations

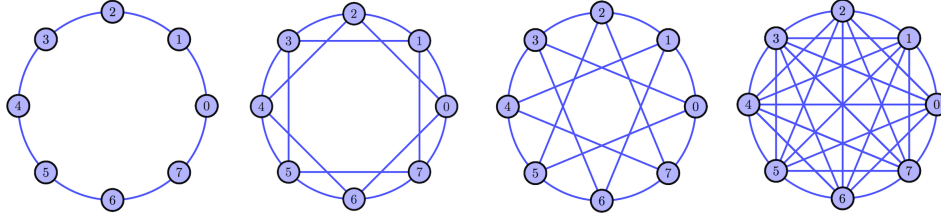


Figure 1: Circulant Graphs with generating sets  $S = \{1\}$ ,  $S = \{1, 2\}$ ,  $S = \{1, 3\}$  and  $S = \{1, 2, 3, 4\}$  (f. left)

We summarize preliminaries in Section 2, and provide an overview of graph e-spline wavelet filterbank constructions in Section 3, before introducing our sampling framework in Section 4. Section 5 features extensions to arbitrary graphs via graph product decompositions, and Section 6 contains concluding remarks with motivations for future directions.

## 2. Preliminaries

A graph  $G = (V, E)$ , with vertex set  $V = \{0, \dots, N-1\}$  of cardinality  $|V| = N$  and edge set  $E$ , is characterized by an adjacency matrix  $\mathbf{A}$ , with  $A_{i,j} > 0$  if vertices  $i$  and  $j$  are adjacent, and  $A_{i,j} = 0$  otherwise, and its degree matrix  $\mathbf{D}$ , which is diagonal with entries  $D_{i,i} = \sum_j A_{i,j}$ . The combinatorial graph Laplacian  $\mathbf{L} = \mathbf{D} - \mathbf{A}$  of undirected graph  $G$  is a positive semi-definite matrix, with a complete set of orthonormal eigenvectors  $\{\mathbf{u}_l\}_{l=0}^{N-1}$  and associated non-negative eigenvalues  $0 = \lambda_0 < \lambda_1 \leq \dots \leq \lambda_{N-1}$ , termed ‘graph-frequencies’. We consider graphs that are undirected, connected, (un-)weighted, and do not allow self-loops; our primary focus however lies on the class of circulant graphs due to their LSI (Linear Shift Invariance) property and regularity which facilitate a more intuitive application and extension of traditional signal processing concepts to the graph domain (as can be seen in Fig. 1). A circulant graph  $G$  with generating set  $S = \{s_1, \dots, s_M\}$ , with  $0 < s_k \leq N/2$ , has adjacency relations between node pairs  $(i, (i \pm s_k)_N), \forall s_k \in S$ , for  $\text{mod}N$  operator  $(\cdot)_N$ , or alternatively, a graph is circulant under some node labelling if its associated graph Laplacian is a circulant matrix [16]. Further, the symmetric, circulant graph Laplacian matrix  $\mathbf{L}$ , with first row  $[l_0 \dots l_{N-1}]$ , has representer polynomial  $l(z) = \sum_{i=0}^{N-1} l_i z^i$  with  $z^j = z^{N-j}$ . The  $2B$ -regular ring lattice  $G$ , within a special sub-class of circulant graphs, has the generating set  $S = \{1, \dots, B\}$ , such that there is an edge between nodes  $i$  and  $j$ , if  $(i - j)_N \leq B$ , and  $\mathbf{L}$  is banded of bandwidth  $B$ .

In this work, we consider graph signals  $\mathbf{x}$  residing on the vertices of a graph  $G$  that are complex-valued, with sample value  $x(i)$  at node  $i$  and represented as the vector  $\mathbf{x} \in \mathbb{C}^N$  [1], while maintaining real weights between connections on  $G$ . The Graph Fourier Transform (GFT)  $\hat{\mathbf{x}}$  of  $\mathbf{x}$  defined on  $G$ , is the expansion in terms of the graph Laplacian eigenbasis  $\mathbf{U} = [\mathbf{u}_0 | \dots | \mathbf{u}_{N-1}]$  such that  $\hat{\mathbf{x}} = \mathbf{U}^H \mathbf{x}$ , where  $H$  denotes the Hermitian transpose, extending the concept of the Fourier transform to the graph domain [1]. Notably, the GFT of circulant graphs can be expressed as a permutation of the DFT-matrix.

A graph (wavelet) filter  $\mathbf{H}$  in the vertex domain generally describes a linear transform

which takes weighted averages (differences) of components of the input signal  $\mathbf{x}$  at a vertex  $i$  within its  $k$ -hop local neighborhood  $N(i, k)$ , and can, where applicable, be expressed as a polynomial in the adjacency matrix  $\mathbf{H} = \sum_{k=0}^{N-1} h_k \mathbf{A}^k$  for suitable  $h_k$  [1]. In [16], a variety of SP concepts and operations on circulant graphs are discussed, including different options to conduct downsampling on such in the context of graph wavelet analysis. Hereby, a given graph signal on  $G$  can be downsampled by 2 with respect to any element  $s_k \in S$ ; for simplicity we resort to the standard downsampling operation with respect to the outmost cycle ( $s_1 = 1$ ) of a given circulant  $G$ , i.e. we skip every other labelled node, assuming that  $G$  is connected such that  $s_1 \in S$ , and  $N = 2^n$  for  $n \in \mathbb{N}$ . After downsampling, the retained nodes can be reconnected to form a coarsened graph for which a number of schemes have been proposed ([29],[36]); we primarily resort to opting for the sparsest possible graph-reconnection under the preservation of circularity, by either retaining the same generating set of the original circulant graph, or alternatively, reducing connectivity by only maintaining existing connections without reconnecting nodes (with exception of preserving  $s_1 \in S$ ). In particular, this ensures that for an originally banded graph adjacency matrix, the bandwidth after coarsening is not increased, and thus facilitates a sparse representation via our graph wavelet filter design. As we will proceed to demonstrate in Sect. 4, the former approach preserves basic graph properties and will be leveraged in our sparse sampling scheme.

### 3. E-Spline Wavelet Analysis on Circulant Graphs

Before we can formulate a framework for graph signal sampling, we need to state the theory of graph spline wavelets and their basic properties, which serve as a crucial element in our interpretation and analysis of sparsity on graphs. Inspired by the circulant graph wavelet filterbank introduced in ([6],[17]) and in light of further detected properties pertaining to the circulant graph Laplacian matrix [18] and its parameterised extension, the e-graph Laplacian [19], which we will briefly state below, we have developed novel families of graph (e-)spline wavelet transforms which form a graph-analogy to the traditional (e-)spline and associated wavelet families. For a more thorough discussion of the comparison and proofs of the accompanying claims, we refer the interested reader to [20].

#### 3.1. Vanishing Moments of the Graph Laplacian

In our ensuing discussion, we distinguish between two main classes of compressible graph signals residing on the vertices of a graph  $G$ :

*(Piecewise) Polynomial:* A graph signal  $\mathbf{p} \in \mathbb{R}^N$  defined on the vertices of a graph  $G$  is (piecewise) polynomial if its labelled sequence of sample values, with value  $p(i)$  at node  $i$ , is the discrete, vectorized version of a standard (piecewise) polynomial, such that  $\mathbf{p} = \sum_{j=1}^K \mathbf{p}_j \mathbf{1}_{[t_j, t_{j+1}[}$ , whereby  $t_1 = 0$  and  $t_{K+1} = N$ , with pieces  $p_j(t) = \sum_{d=0}^D a_{d,j} t^d$ ,  $j = 1, \dots, K$ , for  $t \in \mathbb{Z}_{\geq 0}$ , coefficients  $a_{d,j} \in \mathbb{R}$ , and maximum degree  $D = \deg(p_j(t))$ .

*Complex exponential polynomial:* A complex exponential polynomial graph signal  $\mathbf{x} \in \mathbb{C}^N$  with parameter  $\alpha \in \mathbb{R}$ , is defined such that node  $j$  has sample value  $x(j) = p(j)e^{i\alpha j}$ , for polynomial  $\mathbf{p} \in \mathbb{R}^N$  of degree  $\deg(p(t))$ .

Prior analysis of the graph Laplacian matrix has yielded a distinct annihilation property for the symmetric circulant case. Hereby we adopt the traditional definition of the vanishing moments of order  $N$  of a high-pass filter  $\mathbf{h}$  with taps  $h_k$  as orthogonality of the former with respect to subspaces of polynomials of up to degree  $N - 1$ , i.e. the  $n$ -th order moments  $m_n = \sum_{k \in \mathbb{Z}} h_k k^n$  of  $\mathbf{h}$ , for  $n = 0, \dots, N - 1$  are zero, in order to capture the following results:

**Lemma 3.1.** *For an undirected, circulant graph  $G = (V, E)$  of dimension  $N$ , the associated representer polynomial  $l(z) = l_0 + \sum_{i=1}^B l_i(z^i + z^{-i})$  of graph Laplacian matrix  $\mathbf{L}$ , with first row  $[l_0 \ l_1 \ l_2 \ \dots \ l_2 \ l_1]$ , has two vanishing moments. Therefore, the operator  $\mathbf{L}$  annihilates polynomial graph signals of up to degree  $D = 1$ , subject to a border effect determined by the bandwidth  $B$  of  $\mathbf{L}$ , whereby  $2B < N$ .*

Further, we define the e-graph Laplacian matrix for undirected, circulant graph  $G$  with adjacency matrix  $\mathbf{A}$  and degree  $d = \sum_{j=1}^B 2d_j$  per node with symmetric weights  $d_j = A_{i, (j+i)_N}$ , as  $\tilde{\mathbf{L}}_\alpha = \tilde{\mathbf{D}}_\alpha - \mathbf{A}$ , where  $\tilde{d}_\alpha = \sum_{j=1}^B 2d_j \cos(\alpha j)$  is the parameterised, exponential degree with  $|\tilde{d}_\alpha| \leq d$  and  $\alpha \in \mathbb{R}$ . This operator can be considered a generalization of the classical graph Laplacian, whereby  $\tilde{\mathbf{L}}_\alpha = \mathbf{L}$  for  $\alpha = 0$ , and, although not a positive semi-definite matrix for  $\alpha \neq 0$ , it is of primary interest as a differencing graph operator, as the following property demonstrates:

**Lemma 3.2.** *For an undirected, circulant graph  $G = (V, E)$  of dimension  $N$ , the associated representer polynomial  $\tilde{l}_\alpha(z) = \tilde{l}_0 + \sum_{i=1}^B \tilde{l}_i(z^i + z^{-i})$  of the e-graph Laplacian matrix  $\tilde{\mathbf{L}}_\alpha$ , with first row  $[\tilde{l}_0 \ \tilde{l}_1 \ \tilde{l}_2 \ \dots \ \tilde{l}_2 \ \tilde{l}_1]$ , has two vanishing exponential moments, i.e. the operator  $\tilde{\mathbf{L}}_\alpha$  annihilates complex exponential polynomial graph signals with exponent  $\pm i\alpha$  and  $\deg(p(t)) = 0$ . Unless  $\alpha = \frac{2\pi k}{N}$  for  $k \in [0, N - 1]$ , this is subject to a border effect determined by the bandwidth  $B$  of  $\tilde{\mathbf{L}}_\alpha$ , whereby  $2B < N$ .*

Proofs of the preceding Lemmata entail the detection of roots  $z_\pm = e^{\pm i\alpha}$  for  $\tilde{l}_\alpha(z)$  (with double root  $z = 1$  for  $\tilde{l}_0(z) = l(z)$ ), which indicate two vanishing exponential moments.

### 3.2. Families of Graph E-Spline Wavelets

By leveraging the high-pass filter properties of the e-graph Laplacian operator, we design higher-order critically-sampled and vertex-domain localized graph wavelet filterbanks, which extend classical (e-)spline properties to the graph domain, and distinguish between graph (B-)spline and graph e-spline wavelets respectively:

**Theorem 3.1.** *Given the undirected, and connected circulant graph  $G = (V, E)$  of dimension  $N$ , with adjacency matrix  $\mathbf{A}$  and degree  $d$  per node, we define the higher-order graph-spline wavelet transform (HGSWT), composed of the low-and high-pass filters*

$$\mathbf{H}_{LP} = \frac{1}{2^k} \left( \mathbf{I}_N + \frac{\mathbf{A}}{d} \right)^k \quad (1)$$

$$\mathbf{H}_{HP} = \frac{1}{2^k} \left( \mathbf{I}_N - \frac{\mathbf{A}}{d} \right)^k \quad (2)$$

whose associated high-pass representer polynomial  $H_{HP}(z)$  has  $2k$  vanishing moments. This filterbank is invertible for any downsampling pattern, as long as at least one node retains the low-pass component, while the complementary set of nodes retains the high-pass components.

**Theorem 3.2.** *The higher-order graph e-spline wavelet transform (HGESWT) on a connected, undirected circulant graph  $G$ , is composed of the low-and high-pass filters*

$$\mathbf{H}_{LP_{\vec{\alpha}}} = \prod_{n=1}^T \frac{1}{2^k} \left( \beta_n \mathbf{I}_N + \frac{\mathbf{A}}{d} \right)^k \quad (3)$$

$$\mathbf{H}_{HP_{\vec{\alpha}}} = \prod_{n=1}^T \frac{1}{2^k} \left( \beta_n \mathbf{I}_N - \frac{\mathbf{A}}{d} \right)^k \quad (4)$$

whereby  $\mathbf{A}$  is the adjacency matrix,  $d$  the degree per node and parameter  $\beta_n$  is given by  $\beta_n = \frac{\tilde{d}_{\alpha_n}}{d}$  with  $\tilde{d}_{\alpha_n} = \sum_{j=1}^B 2d_j \cos(\alpha_n j)$  and  $\vec{\alpha} = (\alpha_1, \dots, \alpha_T)$ . Then the high-pass filter annihilates complex exponential polynomials (of  $\deg(p(t)) \leq k-1$ ) with exponent  $\pm i\alpha_n$  for  $n = 1, \dots, T$ . The transform is invertible for any downsampling pattern as long as the eigenvalues  $\gamma_i$  of  $\frac{\mathbf{A}}{d}$  satisfy  $|\beta_n| \neq |\gamma_i|$ ,  $i = 0, \dots, N-1$ , under either of the sufficient conditions

(i)  $k \in 2\mathbb{N}$ , or

(ii)  $k \in \mathbb{N}$  and  $\beta_n, T$  are such that  $\forall \gamma_i, f(\gamma_i) = \prod_{n=1}^T (\beta_n^2 - \gamma_i^2)^k > 0$  or  $f(\gamma_i) < 0$ .

If parameters  $\beta_n$  are such that  $\beta_n = \gamma_i$ , for up to  $T$  distinct values, the filterbank continues to be invertible under the above as long as  $\beta_n \neq 0$  and at least  $\sum_{i=1}^T m_i$  low-pass components are retained at nodes in set  $V_{\alpha}$  such that  $\{\mathbf{v}_{+i,k}(V_{\alpha})\}_{i=1,k=1}^{i=T,k=m_i}$  (and, if  $-\gamma_i$  exists, complement  $\{\mathbf{v}_{-i,k}(V_{\alpha}^C)\}_{i=1,k=1}^{i=T,k=m_i}$ ) form linearly independent sets, where  $m_i$  is the multiplicity of  $\gamma_i$  and  $\{\mathbf{v}_{\pm i,k}\}_{k=1}^{m_i}$  are the eigenvectors respectively associated with  $\pm \gamma_i$ .

In general, we can iterate on the low-pass branch of either transform to obtain a multilevel representation, however, as a consequence of the non-stationarity of the latter (see [37] for the traditional case), modifications to parameters  $\vec{\alpha}$  apply; hereby, we require the parameterization by  $2^j \vec{\alpha}$  at level  $j$  in order to preserve annihilation properties.<sup>1</sup>

The aforementioned transforms can be applied on any undirected circulant graph  $G$ , yet, we observe some noteworthy property distinctions between bipartite and non-bipartite circulant graph cases, as well as when  $|\beta_n| = |\gamma_i|$  is satisfied for some  $n$  and  $i \in [0 \ N-1]$ .

---

<sup>1</sup>Technically, one may describe both graph wavelet filterbank constructions in Thms. 3.1 and 3.2 as ‘non-stationary’ in the sense that the representer polynomials of the respective graph filters at different levels are not necessarily dilates of one another, as a result of their dependence on the adjacency matrix. In particular, unless the coarsened graph, on which the downsampled low-pass output is defined, bears identical edge relations to the initial graph (e.g. when the generating sets are identical for  $2B < N$ ), the representer functions will change with the graph. Nevertheless, the general structure of the filters as polynomials in the adjacency matrix only changes in Thm. 3.2 due to the parameterization by  $\{\beta_i\}_i$ .

### 3.2.1. Properties and Special Cases

It has been established that a scaling function  $H_j(z)$  at level  $j$  can reproduce a function of the form  $P(t)e^{\gamma_m t}$ , where  $\deg P(t) \leq (L_m - 1)$  for multiplicity  $L_m$  of  $\gamma_m$ , if and only if the former is divisible by the term  $R_{2^j \vec{\gamma}}(z)$ ,  $\forall j \leq j_0 - 1$ , where  $R_{\vec{\gamma}}(z) = \prod_{n=1}^M (1 + e^{\gamma_n} z^{-1})$ , with  $\vec{\gamma} = (\gamma_1, \dots, \gamma_M)^T \in \mathbb{C}^M$ , and  $H_j(z)$  has no roots of opposite sign, i.e.  $H_j(z)$  satisfies the generalized Strang-Fix conditions for suitable  $\vec{\gamma}$  ([37], Thm. 1).

Complementing Lemmata 3.1-3.2, we can therefore further deduce that for a circulant and bipartite graph, characterized by all-odd elements  $s_k$  in generating set  $S$ , the low-pass filters  $\mathbf{H}_{LP}$  and  $\mathbf{H}_{LP_{\vec{\alpha}}}$  in eqs. (1) and (3) respectively reproduce (higher-order) polynomial and complex exponential polynomial graph signals, equivalently subject to a border effect that depends on the bandwidth  $Bk$  of the filter matrices, provided  $2Bk < N$ , as is shown in [20].

In addition, one can explicitly describe the Riesz bounds, and hence, the 2-norm condition number of the  $HGESWT$  for a bipartite circulant graph in terms of its spectrum and the given parameters:

**Corollary 3.1.** *The condition number  $C$  of the  $HGESWT$  matrix  $\mathbf{W}$  for a bipartite circulant graph, with downsampling conducted w.r.t.  $s = 1 \in S$ , can be expressed as  $C = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$ , where  $\lambda = \frac{1}{2} \left( \prod_{n=1}^T \frac{1}{2^k} (\beta_n + \gamma)^{2k} + \prod_{n=1}^T \frac{1}{2^k} (\beta_n - \gamma)^{2k} \right)$  for eigenvalues  $\gamma$  of  $\frac{\mathbf{A}}{d}$ .*

*Proof.* See Appendix A1.

For  $\alpha$  of the form  $\frac{2\pi k}{N}$ ,  $k \in [0 \ N - 1]$ , the complex exponential graph signal  $\mathbf{x}$  with values  $x(j) = e^{\pm i\alpha j}$  is periodic and can be perfectly annihilated, and, in the bipartite case, also reproduced by suitable graph filters with e-degree  $\tilde{d}_{\alpha_k} = \sum_{j=1}^B 2d_j \cos\left(\frac{2\pi k j}{N}\right)$  without border effects. In particular, the eigenvalues of normalized adjacency matrix  $\frac{\mathbf{A}}{d}$  can be expressed as  $\gamma_k = \sum_{j=1}^B 2d_j \cos\left(\frac{2\pi k j}{N}\right) / d$ , thereby satisfying  $\gamma_k = \frac{\tilde{d}_{\alpha_k}}{d}$ . Since  $\mathbf{x}$  is an eigenvector of  $\frac{\mathbf{A}}{d}$  corresponding to eigenvalue  $\gamma_k$ , the e-graph Laplacian matrix  $\mathbf{L}_{\alpha}$  can be reinterpreted as a shifted version of  $\mathbf{L}$  whose nullspace consists of (a subset of) its eigenvectors, in the alternative (normalized) representation  $(\gamma_k \mathbf{I}_N - \frac{\mathbf{A}}{d}) \mathbf{x} = \mathbf{0}_N$ . The transform defined in Thm 3.2 continues to be invertible if  $|\beta_n| = |\gamma_i|$ ,  $i \in [0 \ N - 1]$ , is satisfied, up to certain exceptions for  $\beta_n$  (and hence  $\alpha_n$ ), i.a. for the standard alternating downsampling pattern with respect to  $s = 1 \in S$ . The following corollaries capture such restrictions on the invertibility of the graph e-spline wavelet transform for certain  $\alpha_n$ :

**Corollary 3.2.** *The  $HGESWT$  ceases to be invertible for any downsampling pattern as well as fails to reproduce certain graph signals when  $\exists \alpha_i, \alpha_j$  in  $\vec{\alpha}$  such that  $\tilde{d}_{\alpha_i} = -\tilde{d}_{\alpha_j}$  for  $\tilde{d}_{\alpha_l}$  of the form  $\sum_{k=1}^B 2d_k \cos\left(\frac{2\pi k l}{N}\right)$  for  $l \in [0 \ N - 1]$  and  $2B < N$ , including the case  $\tilde{d}_{\alpha_i} = 0$ .*

*Proof.* In the case of a general circulant graph, we have, for  $\alpha_i, \alpha_j$  as above,  $H_{LP_{\alpha_i}}(z) = -H_{HP_{\alpha_j}}(z)$ , or  $\mathbf{H}_{LP_{\alpha_i}} = -\mathbf{H}_{HP_{\alpha_j}}$  in matrix form and vice versa, which leads to annihilation in the low-pass and reproduction in the high-pass branch.



Further, one cannot demonstrate linear independence of the eigenvectors associated with  $\gamma_i, \gamma_j$ , and hence invertibility of the *HGESWT*, for any downsampling pattern, as a consequence of Thm 3.2; for brevity we refer to the complete proof in (Appendix A.2 ,[20]). When  $\tilde{d}_{\alpha_i} = 0$ , the filterbank reduces to the normalized adjacency matrix  $\frac{\mathbf{A}}{d}$  up to a sign per row (and its powers), which is singular if  $\mathbf{A}$  is singular, while its representer polynomial contains the zero root.  $\square$

**Corollary 3.3.** *Let  $\gamma^{DFT} = \{\gamma_i\}_i$  denote the DFT-ordered spectrum of  $\frac{\mathbf{A}}{d}$  for  $\frac{\mathbf{A}}{d} = \mathbf{V}\mathbf{V}^H$  with  $\mathbf{V}$  as the  $N \times N$  DFT-matrix. When downsampling is conducted with respect to  $s = 1 \in S$ , the *HGESWT* ceases to be invertible if  $\exists \alpha_i, \alpha_j$  in  $\vec{\alpha}$  for  $\frac{\tilde{d}_{\alpha_i}}{d}, \frac{\tilde{d}_{\alpha_j}}{d} \in \gamma^{DFT}$ , with respective multiplicities at positions  $M_i = \{i_k\}_k$  and  $M_j = \{j_k\}_k$  in  $\gamma^{DFT}$ , and such that  $\frac{\tilde{d}_{\alpha_i}}{d} = \gamma_i$  is located at position  $(s + N/2)_N, s \in M_j \cup M_i$  of the DFT-ordered spectrum (and vice versa for  $\tilde{d}_{\alpha_j}$ ). When the graph is bipartite, this condition becomes equivalent to that of Cor. 3.2 for the fixed downsampling pattern.*

*Proof.* Given parameters of the form  $\beta_i = \frac{\tilde{d}_{\alpha_i}}{d}$  which are contained in the spectrum  $\gamma^{DFT}$  of  $\frac{\mathbf{A}}{d}$ , we distinguish between the eigenvalues  $\gamma_i$  (and if existent,  $-\gamma_i$ ) and their multiplicities, such that  $|\beta_i| = |\gamma_i|$ , with corresponding eigenvectors  $\mathbf{V}_{\pm\gamma_i} = \{\mathbf{v}_{\pm i,l}\}_l$ . Then the invertibility of the *HGESWT* is conditional upon the eigenvector sets  $\mathbf{V}_{\pm\gamma_i}$  respectively being linearly independent after downsampling each vector by 2 to give  $\mathbf{v}_{+i,l}(V_\alpha)$ , with  $V_\alpha = \{0 : 2 : N - 2\}$  as the set retained nodes (for the detailed proof see [20]). Since  $\mathbf{V}$  is the DFT-matrix and  $\mathbf{V}(V_\alpha, 0 : N - 1) = [\tilde{\mathbf{V}} \ \tilde{\mathbf{V}}]$  with  $\tilde{\mathbf{V}}$  as the DFT of dimension  $N/2$  (up to a normalization constant), we observe that eigenvector pairs  $(\mathbf{v}_k, \mathbf{v}_{k+N/2})$ , at position  $k \in [0 : N - 1]$  become linearly dependent. We therefore need to ensure that the parameters  $\{\beta_i\}_i$  with  $|\beta_i| = |\gamma_i|$  are chosen such that the corresponding values of  $\{\gamma_i\}_i$  (and multiplicities) respectively do not take the aforementioned positions in the DFT-ordered spectrum; for existing  $-\gamma_i$ , the same relation holds for complement  $\mathbf{V}_{-\gamma_i}(V_\alpha^C) = \{\mathbf{v}_{-i,l}(V_\alpha^C)\}_l$ . When the graph is additionally bipartite, we note that given  $\alpha_i, \alpha_j$  at respective positions  $i, j$ , with  $j = (i + N/2)_N$ , due to the relation  $\cos\left(\frac{2\pi k(i+N/2)}{N}\right) = -\cos\left(\frac{2\pi ki}{N}\right)$  for odd  $k$ , we have  $\tilde{d}_{\alpha_i} = -\tilde{d}_{\alpha_j}$  and Cor. 3.2 applies.  $\square$

### 3.2.2. Complementary Graph E-Spline Wavelets for Non-Bipartite Circulants

For non-bipartite circulant graphs, we resort to traditional spectral factorization techniques to create vertex-localized low-pass filters with the required reproduction properties while maintaining the high-pass filter as is; these novel transforms are composed of well-defined analysis and synthesis filters of compact support and can be related to the previous filterbanks via a symmetric circulant transformation filter  $\mathbf{C}$ , depending on the invertibility of the low-pass filters in eqs. (1) and (3) [20].

In particular, given analysis high-pass filter  $H_{HP_{\vec{\alpha}}}(z) = \prod_{n=1}^T \frac{\tilde{l}_{\alpha_n}(z)^k}{(2d)^k}$  with parameterization  $\vec{\alpha} = (\alpha_1, \dots, \alpha_T)$ , we define synthesis low-pass filter  $\tilde{H}_{LP_{\vec{\alpha}}}(z) = H_{HP_{\vec{\alpha}}}(-z)$ , and derive analysis low-pass filter  $H_{LP_{\vec{\alpha}}}(z)$  from

$$P(z) = H_{LP_{\vec{\alpha}}}(z)\tilde{H}_{LP_{\vec{\alpha}}}(z), \quad \text{subject to } P(z) + P(-z) = 2. \quad (5)$$

Further constraints for reproduction properties are imposed in the form of (e-)spline factors in  $H_{LP_{\vec{\alpha}}}(z) = \prod_{n=1}^T (z + 2\cos(\alpha_n) + z^{-1})^k R(z)$ , where  $R(z)$  is determined to

satisfy (5).

**Theorem 3.3.** *Given the undirected, and connected circulant graph  $G = (V, E)$  of dimension  $N$ , with adjacency matrix  $\mathbf{A}$  and degree  $d$  per node, we define the higher-order ‘complementary’ graph e-spline wavelet transform (HCGESWT) via the set of analysis filters:*

$$\mathbf{H}_{LP_{\vec{\alpha}}, an} \stackrel{(*)}{=} \mathbf{C} \bar{\mathbf{H}}_{LP_{\vec{\alpha}}} = \mathbf{C} \prod_{n=1}^T \frac{1}{2^k} \left( \beta_n \mathbf{I}_N + \frac{\mathbf{A}}{d} \right)^k \quad (6)$$

$$\mathbf{H}_{HP_{\vec{\alpha}}, an} = \prod_{n=1}^T \frac{1}{2^k} \left( \beta_n \mathbf{I}_N - \frac{\mathbf{A}}{d} \right)^k \quad (7)$$

and the set of synthesis filters:

$$\mathbf{H}_{LP_{\vec{\alpha}}, syn} = c_1 \mathbf{H}_{HP_{\vec{\alpha}}, an} \circ I_{HP} \quad (8)$$

$$\mathbf{H}_{HP_{\vec{\alpha}}, syn} = c_2 \mathbf{H}_{LP_{\vec{\alpha}}, an} \circ I_{LP} \quad (9)$$

where  $\mathbf{H}_{LP_{\vec{\alpha}}, an}$  is the solution to the system from eq. (5) for  $\vec{\alpha}$  under specified constraints, with coefficient matrix  $\mathbf{C} = \mathbf{H}_{LP_{\vec{\alpha}}, an} \bar{\mathbf{H}}_{LP_{\vec{\alpha}}}^{-1}$  in (\*) arising where applicable (see Cor. 3.3 [20]). Hereby,  $c_i, i \in \{1, 2\}$  are normalization coefficients,  $I_{LP/HP}$  are circulant indicator matrices with first row of the form  $[1 \ -1 \ 1 \ -1 \ \dots]$ , and  $\circ$  denotes the Hadamard product.

The existence of a suitable low-pass filter  $\mathbf{H}_{LP_{\vec{\alpha}}, an}$  to complete the above filterbank is conditional upon the representer polynomial  $H_{HP_{\vec{\alpha}}, an}(z)$  of  $\mathbf{H}_{HP_{\vec{\alpha}}, an}$  yielding no opposing or zero roots for given  $\vec{\alpha}$  so as to satisfy Bézout’s Theorem ([20], [37]). At  $\alpha = 0$ , this gives rise to a regular graph spline wavelet filterbank. We note that while this design is applicable to bipartite circulant graphs, it is less relevant, as the transform given in Thm. 3.2 already provides the desired reproduction properties in that case. As an interesting aside, given a bipartite circulant graph, the special case from Cor. 3.2 can be further extended to provide scenarios which violate Bézout’s Thm. and thus the existence of a complementary filterbank construction when downsampling is conducted with respect to  $s = 1 \in S$ ; specifically, when  $\exists \alpha_i, \alpha_j$  in  $\vec{\alpha}$  such that  $\tilde{d}_{\alpha_i} = -\tilde{d}_{\alpha_j}$  for  $\tilde{d}_{\alpha}$ , the representer polynomial of the graph Laplacian product  $\mathbf{H}_{HP_{\alpha_i}} \mathbf{H}_{HP_{\alpha_j}}$  contains opposing roots due to  $H_{HP_{\alpha_j}}(z) = -H_{HP_{\alpha_i}}(-z)$ .

#### 4. Sampling on Circulant Graphs

The process of sampling a continuous time signal  $x(t)$  in the Euclidean domain traditionally comprises filtering with a given  $h(t)$ , followed by a (uniform) sampling step, which creates the samples  $y_n = (x * h)(t)|_{t=nT}$  at sampling rate  $f_s = \frac{1}{T}$  for period  $T$  ([22], see Fig. 2). At its core, sampling theory provides a bridge between continuous-time and discrete-time signals by seeking to identify ideal methods as well as conditions for the perfect recovery of  $x(t)$  from the given  $y_n$ ; this further extends to identifying distinct classes of  $x(t)$  and suitable filters  $h(t)$  which guarantee perfect recovery.

In a broader sense, sampling in discrete-time can be understood as a dimensionality reduction, which is followed by a dimensionality increase (or interpolation) to recover the original signal [38]. In order to formulate a sampling theory in the graph setting,

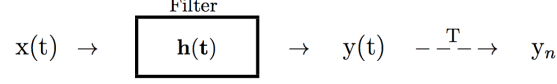


Figure 2: Traditional Sampling Scheme.

additional questions need to be addressed, in particular, on what kind of graph structure the dimensionality-reduced (sampled) signal is defined and how it relates to the original graph, which challenges the classical problem and invites a more sophisticated take on sampling. Sampling theory on graphs can therefore be described as the study of methods and conditions which facilitate the perfect recovery of a graph signal  $\mathbf{x} \in \mathbb{R}^N$  on the vertices of a graph  $G$ , with  $|V| = N$ , from a dimensionality-reduced, possibly graph-filtered signal  $\mathbf{y} \in \mathbb{R}^M$ , extending to the identification of a coarsened graph  $\tilde{G}$ , with  $|\tilde{V}| = M$ , on which  $\mathbf{y}$  is defined. Further intriguing extensions may involve the accompanying recovery of  $G$  from  $\tilde{G}$ , however, in this work we restrict our focus on the former.

#### 4.1. Sparsity and Sampling

Equipped with novel families of sparsifying graph wavelet transforms, we proceed to explore sparse representations on circulant graphs by defining compressible, or ‘wavelet-sparse’ graph signals. Let  $\mathbf{x}_W \in \mathbb{C}^N$  be a signal defined on a circulant graph  $G$  and  $\mathbf{W}_j \in \mathbb{R}^{N/2^j \times N/2^j}$  represent a general graph wavelet transform (GWT) of the form

$$\mathbf{W}_j = \begin{bmatrix} \Psi_{j\downarrow 2} \mathbf{H}_{LP_j} \\ \Phi_{j\downarrow 2} \mathbf{H}_{HP_j} \end{bmatrix},$$

composed of low-and high-pass filters  $\mathbf{H}_{LP_j}, \mathbf{H}_{HP_j} \in \mathbb{R}^{N/2^j \times N/2^j}$ , at level  $j$ , whereby the binary downsampling matrices  $\Psi_{j\downarrow 2}, \Phi_{j\downarrow 2} \in \mathbb{R}^{N/2^{j+1} \times N/2^j}$  sample complementary sets of nodes in the standard alternating pattern with respect to  $s = 1 \in S$ ; hereby even-numbered nodes are retained in the low-pass branch. The multiresolution representation of  $\mathbf{x}_W$ , following iteration on the low-pass branch, then yields

$$\tilde{\mathbf{x}} = \mathbf{W} \mathbf{x}_W = \begin{bmatrix} \mathbf{W}_j & & \\ & \mathbf{I}_{\frac{N(2^j-1)}{2^j}} & \\ & & \ddots \end{bmatrix} \cdots \begin{bmatrix} \mathbf{W}_1 & & \\ & \mathbf{I}_{\frac{N}{2}} & \end{bmatrix} \mathbf{W}_0 \mathbf{x}_W,$$

whereby  $\mathbf{W}$  is the multilevel graph wavelet transform matrix. In order to redefine  $\tilde{\mathbf{x}}$ , whose individual partitions reside on a collective of coarsened graphs, with respect to the original  $G$ , we introduce the permutation matrix  $\mathbf{P}$ , so that, for an appropriate relabelling,  $\mathbf{x} = \mathbf{P} \tilde{\mathbf{x}}$  resides on  $G$ .

Hence, we define the class  $W$  of wavelet- $K$ -sparse graph signals, with  $\mathbf{x}_W \in W$  of dimension  $N$ , through the  $K$ -sparse multiresolution representation  $\mathbf{x} \in \mathbb{C}^N$ ,  $\|\mathbf{x}\|_0 = K$ , via a suitable GWT. In particular, given compressible graph signal  $\mathbf{x}_W$ , we can describe, and hence tailor, the sparsity  $K$  of  $\mathbf{x}$  as a function of decomposition level  $j$  as well as of the bandwidth  $B$  of the graph filter matrix at each  $j$ .

We present results assessing the number of non-zero entries of  $\mathbf{x} \in \mathbb{C}^N$ , for  $\mathbf{x}_W$  in the class

of polynomials, which are generalizable to complex exponential polynomials on graphs. Hereby, we conduct node reconnection such that the coarse graphs retain their original generating set after downsampling at each level, thus maintaining constant filter support, and omit generalizations to other graph coarsening schemes, such as Kron-reduction, for simplicity.

**Corollary 4.1.** *Consider an undirected, circulant graph  $G$  of dimension  $N$  and bandwidth  $\tilde{B}$ , and let  $\mathbf{x}$  be the multiresolution decomposition of graph signal  $\mathbf{x}_W$ , which is a 1-piece polynomial of maximum degree  $D \leq 2k - 1$ , on  $G$  via the  $j$ -level GWT matrix  $\mathbf{W}$ .*

- (i) *Let  $\mathbf{W}$  be the HGSWT of order  $2k$ , whereby the corresponding low-and high-pass graph filter matrices each have bandwidth  $B = k\tilde{B}$ , and assume that  $B$  is sufficiently small such that  $\sum_{n=0}^l \frac{B}{2^n} \leq \frac{N}{2^{l+1}}$  at each level  $l \leq j - 1$ . The resulting  $\mathbf{x} = \mathbf{P}\mathbf{W}\mathbf{x}_W$  is  $K$ -sparse, whereby  $K = \frac{N}{2^j} + B(2(j-1) + 2^{1-j})$ , when  $B = 2^{j-1}r$ ,  $r \in \mathbb{Z}^+$ .*
- (ii) *Let  $\mathbf{W}$  be the HCGSWT of order  $2k$ , whereby the corresponding low-and high-pass graph filter matrices have bandwidth  $T$  and  $B = k\tilde{B}$  respectively, such that  $B + \sum_{n=1}^l \frac{T}{2^n} \leq \frac{N}{2^{l+1}}$  at each level  $l \leq j - 1$ . The resulting  $\mathbf{x}$  is  $K$ -sparse, whereby  $K = \frac{N}{2^j} + Bj + T(j + 2^{1-j} - 2)$ , when  $T = 2^{j-1}r$ ,  $r \in \mathbb{Z}^+$ .*
- (iii) *Let  $\mathbf{W}$  be the HGSWT at  $j = 0$ , with the alternative ‘minimum’ downsampling pattern, which retains only one low-pass component. Then  $\mathbf{x}$  is  $K$ -sparse with  $K = 2B$ .*

*Proof.* See Appendix A2.

When  $B \in \mathbb{Z}^+$ , the results of (i) & (ii) in Cor. 4.1 apply up to a small correction term, which increases with the number of levels  $j$ . For the multiresolution decomposition of periodic (complex exponential) graph signals with parameter  $\alpha = \frac{2\pi k}{N}$ ,  $k \in \mathbb{N}$ , we have the maximum sparsity of  $K = \frac{N}{2^j}$  at  $j$  levels.

#### 4.2. The Graph FRI-framework

The traditional FRI-framework is build on the central result that certain classes of non-bandlimited signals with a finite rate of innovation can be sampled and perfectly reconstructed using kernels of compact support, which satisfy Strang-Fix conditions ([21], [22]); in the discrete domain, this prominently entails that a  $K$ -sparse signal vector  $\mathbf{x} \in \mathbb{R}^N$  can be perfectly reconstructed from  $M \geq 2K$  consecutive sample values  $y_n$  of the measurement vector  $\mathbf{y} = \mathbf{F}\mathbf{x}$ , where  $\mathbf{F} \in \mathbb{C}^{N \times N}$  is the DFT matrix, of the form

$$y_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{K-1} x_{c_k} e^{-i2\pi c_k n/N} = \sum_{k=0}^{K-1} \alpha_k u_k^n \quad (10)$$

with weights  $x_{c_k}$  of  $\mathbf{x}$  at positions  $c_k$ . Hereby, the locations  $u_k = e^{-i2\pi c_k n/N}$  and amplitudes  $\alpha_k = x_{c_k}/\sqrt{N}$  are successively recovered using a reconstruction algorithm known as Prony’s method [22]. In particular, the filtering (or acquisition) of a sparse signal with  $\mathbf{F}$  facilitates its exact reconstruction from a dimensionality-reduced version in the Fourier domain.

The insight that the graph frequency-ordered GFT basis of an arbitrary circulant graph  $G$  can be expressed as the DFT-matrix subject to a graph-dependent permutation of

columns, motivates an extension of sparse sampling to the graph-domain, and in the following, we proceed to formulate the FRI-framework for signals on circulant graphs, which we term the Graph FRI-framework (GFRI).

**Theorem 4.1.** (*Graph-FRI*) Define the permuted GFT basis  $\mathbf{U}$  of undirected circulant graph  $G$  such that  $\mathbf{U}^H$  is the DFT-matrix. We can sample and perfectly reconstruct a (wavelet-)  $K$ -sparse graph signal (with multiresolution)  $\mathbf{x} \in \mathbb{C}^N$ , on the vertices of circulant  $G$  using the dimensionality-reduced GFT representation  $\mathbf{y} = \mathbf{U}_M^H \mathbf{x}$ ,  $\mathbf{y} \in \mathbb{C}^M$ , whereby  $\mathbf{U}_M^H$  are the first  $M$  rows of  $\mathbf{U}^H$ , as long as  $M \geq 2K$ .

*Proof.* See Appendix A.3.

Similarly as in the traditional case, the proof of this theorem follows via application of Prony's method. In reference to our previous sparsity analysis, we therefore require at least  $K < \frac{N}{2}$  for a given graph signal  $\mathbf{x}$ , since  $M \geq 2K$ , for  $N = 2^n, n \in \mathbb{Z}^+$  by initial assumption.

#### 4.2.1. Graph Coarsening for GFRI

The problem of downsampling a signal on a graph  $G = (V, E)$  along with graph coarsening, as the task of determining the reduced set of vertices and edges of the coarsened graph  $\tilde{G} = (\tilde{V}, \tilde{E})$ , are inherent to GSP theory and represent one of the challenges that the complex data dependencies of graphs impose on traditional signal processing. A variety of approaches have been formulated [36], ranging from spectral graph partitioning, where the largest graph Laplacian eigenvector is used to determine a downsampling pattern, up to graph-specific operations such as for bipartite graphs ([1], [7]), which naturally comprise a partitioning into two disjoint sets of nodes. Hereby, reconnection may be conducted to satisfy a range of properties, and is an accompanying problem in itself. In the context of multilevel graph wavelet analysis, the properties of most interest to us are preservation of circularity and a sparse GWT representation, and as implied by our foregoing discussion, the latter is achieved when the bandwidth of the graph Laplacian is small, i.e. minimal reconnection is conducted.

In our current set-up, we are interested in identifying the graph structure associated to the dimensionality reduced GFT-representation  $\mathbf{y}$ , yet, conversely to general graph coarsening approaches, we need to extract an appropriate downsampling pattern as well as a reconnection strategy from the information given by the spectral coefficients at hand, rather than impose a set of desired properties in the first instance. The difficulty is posed by the fact that  $\mathbf{y}$  resides in the graph spectral domain and does not directly give rise to a specific downsampling pattern in the vertex domain.

In the traditional FRI-framework [22], a given sparse signal can be sampled with a general exponential reproducing kernel  $\phi(t)$ , not restricted to the complex exponentials of the DFT as previously shown, whereby the function  $\phi(t)$  in continuous-time and its shifted versions, is such that it can reproduce exponentials for a proper choice of coefficients

$$c_{m,n} \sum_{n \in \mathbb{Z}} c_{m,n} \phi(t - n) = e^{\alpha_m t} \quad \text{for } \alpha_m \in \mathbb{C}, m = 0, \dots, P. \quad (11)$$

We note that the coefficients  $c_{m,n}$  in (11) can be expressed as  $c_{m,n} = c_{m,0} e^{\alpha_m n}$ , where  $c_{m,0} = \int_{-\infty}^{\infty} e^{\alpha_m x} \tilde{\phi}(x) dx$  [22]. Hereby, the functions  $\phi(t)$  and  $\tilde{\phi}(t)$  form a quasi-biorthonormal

set, with biorthonormality as a special case ([39], [40]).

Inspired by this notion of sampling a sparse signal in a multi-layered scheme, we extend our Graph FRI-framework by expressing the reduced GFT-basis  $\mathbf{U}_M^H$  as the product between a fat coefficient matrix  $\mathbf{C}$  and a row-reduced low-pass GWT filter, which can reproduce complex exponential graph signals as per Thms 3.2 and 3.3.

We proceed to demonstrate the feasibility of this scheme by first proving the existence of such a  $\mathbf{C}$  and its relation to a row-reduced DFT-matrix in the following.

**Lemma 4.1.** *Let  $\mathbf{U}_M^H$  be the reduced GFT-basis of undirected circulant graph  $G$ , as defined in Thm. 4.1, and  $\mathbf{E}_{\vec{\alpha}} \in \mathbb{R}^{N \times N}$  a low-pass graph filter matrix in the e-spline GWT family (see Thms. 3.2, 3.3), which can reproduce complex exponential graph signals with parameter  $\vec{\alpha} = (\alpha_0, \dots, \alpha_{M-1}) = \left(0, \dots, \frac{2\pi k}{N}, \dots, \frac{2\pi(M-1)}{N}\right)$ . We thus have  $\mathbf{U}_M^H = \mathbf{C}\Psi_{\downarrow 2}\mathbf{E}_{\vec{\alpha}}$ , whereby  $\Psi_{\downarrow 2} \in \mathbb{R}^{N/2 \times N}$  is a binary sampling matrix which retains even-numbered nodes, and  $\mathbf{C} \in \mathbb{C}^{M \times N/2}$  is a coefficient matrix. Further,  $\mathbf{C} = \hat{\mathbf{C}}\tilde{\mathbf{U}}_M^H$ , whereby  $\hat{\mathbf{C}} \in \mathbb{C}^{M \times M}$  is diagonal and  $\tilde{\mathbf{U}}^H$  is the DFT matrix of dimension  $N/2$ .*

*Proof.* Consider the general complementary graph e-spline wavelet filterbank (Thm. 3.3) with respective analysis and synthesis matrices

$$\mathbf{W} = \begin{bmatrix} \Psi_{\downarrow 2}\mathbf{H}_{LP\alpha} \\ \Phi_{\downarrow 2}\mathbf{H}_{HP\alpha} \end{bmatrix}, \quad \tilde{\mathbf{W}} = \begin{bmatrix} \Psi_{\downarrow 2}\tilde{\mathbf{H}}_{LP\alpha} \\ \Phi_{\downarrow 2}\tilde{\mathbf{H}}_{HP\alpha} \end{bmatrix}$$

such that  $\tilde{\mathbf{W}}^T\mathbf{W} = \mathbf{I}_N$ , whereby the high-pass representer polynomials at both branches possess the same number of vanishing moments, i.e. roots at  $z = e^{\pm i\alpha}$  for some  $\alpha \in \mathbb{R}$ . Let  $\mathbf{x} \in \mathbb{C}^N$  be a complex exponential graph signal of the form

$$\mathbf{x} = [e^{i\alpha 0} \quad e^{i\alpha 1} \quad e^{i\alpha 2} \quad \dots \quad e^{i\alpha(N-1)}]^T$$

whereby  $\alpha = -\frac{2\pi k}{N}$ , i.e.  $\mathbf{x}^T$  is the  $(k+1)$ -th row of the (unnormalized) DFT-matrix, and define  $\mathbf{y} = \mathbf{H}_{LP\alpha}\mathbf{x} = c\mathbf{x}$  for  $c \in \mathbb{R}$  (also an eigenvalue of  $\mathbf{H}_{LP\alpha}$ ), such that

$$\Psi_{\downarrow 2}\mathbf{H}_{LP\alpha}\mathbf{x} = c\Psi_{\downarrow 2}\mathbf{x} = c[e^{i\alpha 0} \quad e^{i\alpha 2} \quad e^{i\alpha 4} \quad \dots \quad e^{i\alpha(N-2)}]^T = \mathbf{y}(0:2:N-2) = \mathbf{y}_{\downarrow 2}$$

which denotes a scalar multiple of the  $(k+1)$ -th row of the DFT of dimension  $N/2$ , since

$$[e^{i\alpha 0} \quad e^{i\alpha 2} \quad e^{i\alpha 4} \quad \dots \quad e^{i\alpha(N-2)}]^T = [e^{i(2\alpha)0} \quad e^{i(2\alpha)1} \quad e^{i(2\alpha)2} \quad \dots \quad e^{i(2\alpha)(N/2-1)}]^T$$

with  $2\alpha = -\frac{2\pi k}{N/2}$ . We obtain

$$[(\Psi_{\downarrow 2}\tilde{\mathbf{H}}_{LP\alpha})^T \quad (\Phi_{\downarrow 2}\tilde{\mathbf{H}}_{HP\alpha})^T] \begin{bmatrix} \Psi_{\downarrow 2}\mathbf{H}_{LP\alpha} \\ \Phi_{\downarrow 2}\mathbf{H}_{HP\alpha} \end{bmatrix} \mathbf{x} = [(\Psi_{\downarrow 2}\tilde{\mathbf{H}}_{LP\alpha})^T \quad (\Phi_{\downarrow 2}\tilde{\mathbf{H}}_{HP\alpha})^T] \begin{bmatrix} \mathbf{y}_{\downarrow 2} \\ \mathbf{0}_{N/2} \end{bmatrix},$$

but since  $(\Phi_{\downarrow 2}\tilde{\mathbf{H}}_{HP\alpha})^T\mathbf{0}_{N/2} = \mathbf{0}_{N/2}$ , neither  $\mathbf{0}_{N/2}$  nor  $(\Phi_{\downarrow 2}\tilde{\mathbf{H}}_{HP\alpha})^T$  contribute, and we can thus write

$$[(\Psi_{\downarrow 2}\tilde{\mathbf{H}}_{LP\alpha})^T \quad (\Phi_{\downarrow 2}\tilde{\mathbf{H}}_{HP\alpha})^T] \begin{bmatrix} \mathbf{y}_{\downarrow 2} \\ \mathbf{0}_{N/2} \end{bmatrix} = [(\Psi_{\downarrow 2}\tilde{\mathbf{H}}_{LP\alpha})^T] [\mathbf{y}_{\downarrow 2}] = \mathbf{x}$$

i.e. linear combinations of the columns of  $(\Psi_{\downarrow 2} \tilde{\mathbf{H}}_{LP_\alpha})^T$  reproduce  $\mathbf{x}$ . Rewriting the former, we obtain  $\mathbf{y}_{\downarrow 2}^T \Psi_{\downarrow 2} \tilde{\mathbf{H}}_{LP_\alpha} = \mathbf{x}^T$ , and reversing the sequence of  $\mathbf{W}$  and  $\tilde{\mathbf{W}}$ , and letting  $\mathbf{E}_\alpha = \mathbf{H}_{LP_\alpha}$ , we arrive at

$$[\mathbf{c}^T] \Psi_{\downarrow 2} \mathbf{E}_\alpha = \mathbf{x}^T$$

with  $\mathbf{c} \in \mathbb{C}^{N/2}$  ( $\mathbf{c} = \hat{c} \mathbf{x}_{\downarrow 2}$  for eigenvalue  $\hat{c}$  of  $\tilde{\mathbf{H}}_{LP_\alpha}$ ),  $\Psi_{\downarrow 2} \mathbf{E}_\alpha \in \mathbb{R}^{N/2 \times N}$ , and  $\mathbf{x}^T \in \mathbb{C}^N$ . By generalizing the RHS to incorporate  $M$  stacked complex exponential vectors  $\mathbf{x}$  to form the transposed DFT-matrix  $(\mathbf{U}_M^H)^T$ , we can similarly show

$$\mathbf{C} \Psi_{\downarrow 2} \mathbf{E}_{\vec{\alpha}} = \mathbf{U}_M^H,$$

with  $\mathbf{C} = \hat{\mathbf{C}} \tilde{\mathbf{U}}_M^H$  and  $\vec{\alpha} = (\alpha_1, \dots, \alpha_M)$ . In particular, the matrix  $\hat{\mathbf{C}}$  is diagonal, while  $\tilde{\mathbf{U}}_M^H$  represents the first  $M$  rows of the DFT-matrix of dimension  $N/2$ .

In the case of a bipartite graph, using Thm. 3.2, we can proceed similarly, and obtain  $[(\Psi_{\downarrow 2} \tilde{\mathbf{H}}_{LP_\alpha})^T] [\mathbf{y}_{\downarrow 2}] = \mathbf{x}$  with  $\mathbf{y}_{\downarrow 2} = c \mathbf{x}_{\downarrow 2}$  for some  $c \in \mathbb{R}$ , for synthesis low-pass filter  $\tilde{\mathbf{H}}_{LP_\alpha}$ , which reveals that it reproduces complex exponentials with parameter  $\pm \alpha$ , just as the analysis low-pass filter, despite not being of the same support. In particular, as we have shown in [20], the inherent biorthogonality constraints of the wavelet transform impose that the representer polynomials of the bipartite HGESWT contain opposing roots respectively for analysis and synthesis, i.e. we have  $H_{LP_\alpha}(z) = -z^{-1} H_{HP_\alpha}(-z)$  and  $\tilde{H}_{LP_\alpha}(z) = -z \tilde{H}_{HP_\alpha}(-z)$ . Thus we may interchange the order of synthesis and analysis branch, and obtain  $[(\Psi_{\downarrow 2} \mathbf{H}_{LP_\alpha})^T] [\hat{c} \mathbf{x}_{\downarrow 2}] = \mathbf{x}$ , confirming our previous result that the columns of the adjacency-matrix based, analysis low-pass filter  $(\Psi_{\downarrow 2} \mathbf{H}_{LP_\alpha})^T$  reproduce  $\mathbf{x}$  as a consequence of the generalized Strang-Fix conditions ([37], [41]).  $\square$

In order to eventually identify a graph coarsening scheme within our sparse sampling framework, we begin by noting that a sensible downsampling pattern can already be extracted from the previous result, namely the low-pass graph filter samples every other node in keeping with the standard circulant downsampling pattern with respect to  $s = 1 \in S$ . A popular graph coarsening scheme known as Kron-reduction [29] employs Schur complementation based on the given node sampling pattern, whereby the graph-Laplacian matrix  $\tilde{\mathbf{L}}$  of the coarsened graph is evaluated from the graph Laplacian matrix  $\mathbf{L}$  of initial graph  $G$  and set  $V_\alpha = \{0 : 2 : N - 2\}$  of retained nodes:

$$\tilde{\mathbf{L}} = \mathbf{L}(V_\alpha, V_\alpha) - \mathbf{L}(V_\alpha, V_\alpha^C) \mathbf{L}(V_\alpha^C, V_\alpha^C)^{-1} \mathbf{L}(V_\alpha, V_\alpha^C)^T.$$

In particular, it can be shown that for  $V_\alpha$  as set above and symmetric circulant  $\mathbf{L}$ , the resulting coarsened graph will preserve these properties ([29], [17]); a drawback is however that it leads to denser graphs, i.e. for a banded circulant matrix, the resulting lower-dimensional Schur complement will be of equal or larger bandwidth which proves destructive in our sparsity-driven filterbank construction and analysis.

Alternatively, we propose to conduct sampling in the spectral domain of the graph at hand by leveraging the fact that any graph Laplacian eigenvector  $\mathbf{u}_k$  of  $G$  can be interpreted as a graph signal on its vertices with sample value  $u_k(i)$  at node  $i$ , suggesting that an extracted downsampling pattern may be equivalently applied to the eigenbasis of  $G$ .

**Lemma 4.2.** Consider an undirected circulant graph  $G$  with generating set  $S$ , and adjacency matrix  $\mathbf{A} = \frac{1}{N}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \in \mathbb{R}^{N \times N}$  with bandwidth  $B$ , where  $\frac{1}{\sqrt{N}}\mathbf{U}^H$  is the DFT matrix. We downsample by 2 via the binary matrix  $\mathbf{\Psi}_{\downarrow 2} \in \mathbb{R}^{N/2 \times N}$  on the first  $N/2$  rows in  $\mathbf{U}^H$  and eigenvalues  $\mathbf{\Lambda}$ , such that  $\tilde{\mathbf{U}}^H = \mathbf{U}_{0:N/2-1}^H \mathbf{\Psi}_{\downarrow 2}^T$ , and  $\tilde{\mathbf{\Lambda}} = \mathbf{\Psi}_{\downarrow 2} \mathbf{\Lambda} \mathbf{\Psi}_{\downarrow 2}^T$ . The resulting adjacency matrix  $\tilde{\mathbf{A}} = \frac{1}{N/2} \tilde{\mathbf{U}} \tilde{\mathbf{\Lambda}} \tilde{\mathbf{U}}^H \in \mathbb{R}^{N/2 \times N/2}$  is circulant with the same generating set  $S$  as  $G$ .

*Proof.* See Appendix A.4.

We observe that the above result is further reinforced by the fact that adjacency and graph Laplacian matrices of regular graphs possess the same basis, which as we will see in a later discussion, is not upheld for e.g. a path graph. In particular, we note that Lemma 4.2 gives rise to an intriguing coarsening strategy for circulant graphs as it preserves both the original connectivity of the graph by retaining the same generating set, as well as its spectral properties given that its eigenvalues and eigenbasis are respectively composed of a subset and subpartition of the original.

Following a generalization of Lemma 4.1, and the preceding discussion, we formulate the graph coarsening scheme to complement our Graph-FRI framework:

**Theorem 4.2.** Given GFT  $\mathbf{y} \in \mathbb{C}^M$ , from Thm 4.1, we determine the coarsened graph  $\tilde{G} = (\tilde{V}, \tilde{E})$  associated with the dimensionality-reduced graph signal  $\tilde{\mathbf{y}} \in \mathbb{C}^{\tilde{M}}$ , via the decomposition

$$\mathbf{y} = \mathbf{U}_M^H \mathbf{x} = \mathbf{C} \prod_{j=0}^{J-1} (\mathbf{\Psi}_{j\downarrow 2} \mathbf{E}_{2^j \vec{\alpha}}) \mathbf{x} = \mathbf{C} \tilde{\mathbf{y}}$$

whereby  $\mathbf{U}_M^H \in \mathbb{C}^{M \times N}$  is the row-reduced permuted GFT basis (DFT-matrix),  $\mathbf{C} \in \mathbb{C}^{M \times \tilde{M}}$  is a coefficient matrix with  $\tilde{M} = \frac{N}{2^J}$  given  $M$ ,  $\mathbf{\Psi}_{j\downarrow 2} \in \mathbb{R}^{N/2^{j+1} \times N/2^j}$  is a binary sampling matrix which retains even-numbered nodes, and  $\mathbf{E}_{2^j \vec{\alpha}} \in \mathbb{R}^{N/2^j \times N/2^j}$  is a (higher-order) graph e-spline low-pass filter on  $\tilde{G}_j$ , which reproduces complex exponentials at level  $j$ , with parameter  $\vec{\alpha} = (\alpha_0, \dots, \alpha_{M-1}) = \left(0, \dots, \frac{2\pi(M-1)}{N}\right)$ . The associated coarsened graphs  $\tilde{G}_j$  at levels  $j \leq J$  can be determined following two different schemes:

- (i) Perform Kron-reduction at each level  $j \leq J$  using the pattern  $\mathbf{V}_\alpha$  in  $\mathbf{\Psi}_{j\downarrow 2}$  to obtain  $\mathbf{L}_j$
- (ii) Define eigenbasis  $(\tilde{\mathbf{U}}_j, \tilde{\mathbf{\Lambda}}_j) \in \mathbb{C}^{N/2^j \times N/2^j}$  at each level  $j \leq J$  through the application of  $\mathbf{\Psi}_{j-1\downarrow 2}$  on  $(\tilde{\mathbf{U}}_{j-1}, \tilde{\mathbf{\Lambda}}_{j-1})$  (see Lemma 4.2). The coarse graph  $\tilde{G}_j$  for graph signal  $\tilde{\mathbf{y}}_j = \prod_{k=0}^{j-1} (\mathbf{\Psi}_{k\downarrow 2} \mathbf{E}_{2^k \vec{\alpha}}) \mathbf{x}$ , has adjacency matrix

$$\mathbf{A}_j = (2^j/N) \tilde{\mathbf{U}}_j \tilde{\mathbf{\Lambda}}_j \tilde{\mathbf{U}}_j^H$$

which preserves the generating set  $S$  of  $G$  for a sufficiently small bandwidth.

Consequentially, the edge set of the coarsened graph associated with the GFRI-framework is not unique, and we have explored two possible approaches which satisfy our connectivity constraints of symmetry and circularity. Kron-reduction preserves basic graph characteristics, yet, while taking into account the entire graph adjacency relations in the



computation of the coarsened version, it provides little general intuition on the topology of the latter. In contrast, the alternative spectral reduction technique, is shown to preserve the original graph connectivity by retaining its generating set (see Fig. 3), thereby simultaneously alleviating the issue of an increasing bandwidth.

We summarize the graph sampling framework, further illustrated in Fig. 4, as the filtering of a sparse graph signal  $\mathbf{x}$  on  $G$  with a graph e-spline low-pass filter followed by dimensionality reduction, giving rise to signal  $\tilde{\mathbf{y}}$  on coarsened  $\tilde{G}$ , which is subsequently transformed into the further dimensionality-reduced, scaled spectral graph domain, resulting in the representation  $\mathbf{y}$ . Graph signals  $\mathbf{x}_W$  with a sparse multiresolution representation via a GWT can be similarly sampled following an initial sparsification step.

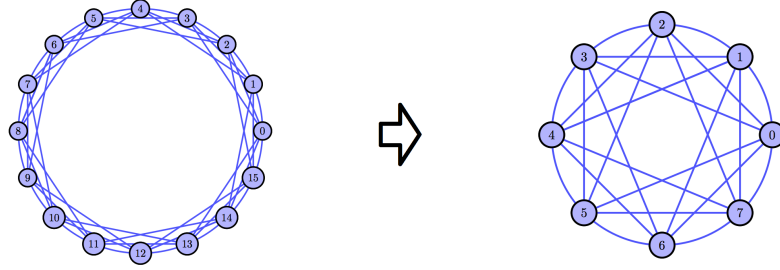


Figure 3: Graph Coarsening for a Circulant Graph with  $S = \{1, 2, 3\}$ .

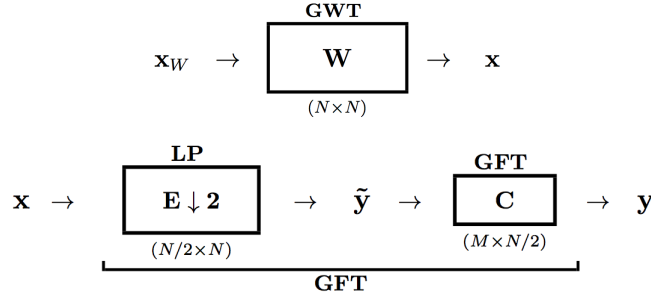


Figure 4: Sampling Scheme with Preceding Sparsification Step and One Level of Coarsening.

The matrix  $\mathbf{E}_{2^j \tilde{\alpha}}$ , representing a higher-order, vertex-localized graph e-spline wavelet low-pass filter parameterised by  $\tilde{\alpha} 2^j$  at level  $j$ , is of the form of the low-pass filter designed in eqns. (3) (Thm. 3.2) or (6) (Thm. 3.3), depending on whether the graph at hand is bipartite or not; we thus implicitly assume that  $\mathbf{E}_{2^j \tilde{\alpha}}$  can reproduce complex exponential graph signals, with  $\tilde{\alpha}$  as specified. As the filter construction in either case is based on the combination/convolution of different graph e-spline basis functions, it should be noted that the resulting higher-order function may for certain scenarios contain opposing roots, which in the non-bipartite case would violate Bézout's Thm [37], and thus the necessary biorthogonality condition for filterbank construction. While for complementary graph

wavelet filterbanks, it is more intricate to deduce generalized conditions on when exactly this occurs, we can further specify conditions under which the HGESWT in the bipartite graph case loses reproduction properties and/or ceases to be invertible, as formulated in Cor. 3.2<sup>2</sup>.

Further to Thm 4.2, we proceed to specify restrictions on the number of samples  $M$  and levels  $J$  to ensure the reproduction via low-pass graph filter  $\mathbf{E}_{\vec{\alpha}}$ . Certain rows of the DFT-matrix cannot be reproduced using a real-valued symmetric graph filter, as alluded to in (Sect. 3, [20]); for instance, parameter  $\alpha = \pm\pi/2$  induces the opposing factors  $(1 - iz)(1 + iz)$  of complex conjugates in the representer polynomial of a parameterized graph Laplacian filter, which violates Bézouts equality for complementary filterbank construction, as well as prevents the reproduction of the corresponding complex exponential in the DFT via a *HGESWT*-based low-pass graph filter. It follows that we cannot reproduce consecutive rows of the DFT-matrix beyond its  $N/4$ -th row.

We therefore need to ensure within the multiresolution analysis that

$$\alpha 2^{j-1} = \frac{2\pi k 2^{j-1}}{N} < \frac{2\pi(N/4)}{N}, \quad \forall j \leq J$$

such that  $k < \frac{N}{2^{j+1}}$  or  $J < \log_2\left(\frac{N}{k}\right) - 1$ , whereby  $J$  is the total number of levels; in other words, we can approximate the DFT-matrix up to its  $M = k + 1$ -th row for a certain number of levels, with parameters  $\tilde{M} = \frac{N}{2^J}$  and  $M = \frac{N}{2^{J+1}} = \frac{\tilde{M}}{2}$  of Thm 4.2. This coincides with the biorthogonality constraint for traditional e-spline wavelets outlined in [43], [37] which ensures that the corresponding filters do not contain opposing roots, whereby given distinct  $\gamma, \gamma' \in \vec{\gamma}$ ,  $2^j(\gamma - \gamma') = i\pi(2k + 1)$  must not be satisfied for some  $k \in \mathbb{Z}$  at level  $j \leq j_0 - 1$ . Depending on the signal set, we may not always satisfy  $K = \frac{N}{2^{j+2}}$  exactly, in which case we simply require  $M < \tilde{M}$  and generalize the formulae accordingly.

Since the graph e-spline filter functions of Thm. 3.2 give rise to the traditional e-spline when the graph at hand is a simple cycle, the aforementioned condition with  $\gamma = i\alpha$  is sufficient in that case; nevertheless as a result of the complex connectivity of circulant graphs, we need to impose further restrictions on  $\vec{\alpha}$ .

For an unweighted bipartite graph with consecutive, odd elements  $s_k \in S$  and even generating set cardinality  $|S|$ , and as a particular case of Cor. 3.2, we additionally observe at  $\alpha = \pm\pi/4$  for e-degree  $\tilde{d}_\alpha = \sum_{k \in \mathbb{Z}} 2 \cos(\alpha(2k + 1)) = 0$ , since  $\cos(\alpha(2k + 1)) = \frac{\sqrt{2}}{2}$ ,  $k = 0, 3, 4, 7, 8, \dots$ , and  $\cos(\alpha(2k + 1)) = -\frac{\sqrt{2}}{2}$ ,  $k = 1, 2, 5, 6, \dots$ ; thus similarly as before, the associated graph filter polynomials contain opposing roots and we can only approximate consecutive rows of the DFT-matrix up to at most the  $N/8$ -th row. This translates into the constraints  $k < \frac{N}{2^{j+2}}$  or  $J < \log_2\left(\frac{N}{k}\right) - 2$ , and  $M = \frac{N}{2^{j+2}} = \frac{\tilde{M}}{4}$ . Hereby,  $M$  may be subjected to further reduction, as a consequence of increasing eigenvalue multiplicities for  $\tilde{d}_\alpha = \frac{\gamma_i}{d} = 0$  for different graph-connectivities.

---

<sup>2</sup>In particular, this amounts to showing when  $\tilde{d}_{\alpha_j} = \sum_{k \in 2\mathbb{Z}+1} 2d_k \cos\left(\frac{2\pi j k}{N}\right) = -\tilde{d}_{\alpha_t} = \sum_{k \in 2\mathbb{Z}+1} 2d_k \cos\left(\frac{2\pi k}{N}\left(t \pm \frac{N}{2}\right)\right)$  can occur, i.e. for which  $\alpha_j = \frac{2\pi j}{N}$ ,  $\alpha_t = \frac{2\pi t}{N}$ , weights  $d_k$  and graph connectivity  $k \in 2\mathbb{Z} + 1$ , the scheme ceases to be valid.

In general, if  $G$  is bipartite, we need to ensure that no parameters  $\alpha_i, \alpha_j \in \vec{\alpha}$  satisfy  $\tilde{d}_{\alpha_i} = -\tilde{d}_{\alpha_j}$  at all levels in order to preserve the invertibility property of Thm 3.2, and thus consider consecutive frequencies (or consecutive rows of the DFT), only up to some cut-off frequency with  $\alpha_k = \frac{2\pi k}{N}$  at position  $k + 1$ , such that for  $i, j \leq k$ ,  $\tilde{d}_{\alpha_i} \neq -\tilde{d}_{\alpha_j}$ . As per Cor. 3.3,  $\tilde{d}_{\alpha_i} = -\tilde{d}_{\alpha_j}$  is satisfied when  $j = (i + N/2)_N$ , which cannot occur for the previously derived constraint  $i, j < N/4$ , however, as a result of large multiplicities (associated with higher graph connectivity) the previous equality may still arise for some  $i, j$ . An example is given by the normalized adjacency matrix of the unweighted complete (circulant) bipartite graph with bipartite sets of equal size  $N/2$ , whose eigenvalues are  $\gamma_{\max/\min} = \pm 1$  of respective multiplicity  $m = 1$ , and  $\gamma_i = 0$  of multiplicity  $N - 2$ . These are necessary conditions for the existence of a suitable low-pass filter via the HGESWT, we note, however, that the set of special cases presented here is not exhaustive. It can be further stated that the condition  $j \neq (i + N/2)_N$  also needs to be satisfied for non-bipartite circulant graphs in the HCGESWT, which equivalently follows from the traditional biorthogonality constraints as well as from a special case of the presented graph spline wavelet transform designs.<sup>3</sup>

#### 4.3. Extensions to Path Graphs

The path graph, which corresponds to a simple cycle without the periodic extension, bears similar properties to its circulant counterpart; it is bipartite and its graph Laplacian eigenvectors can be represented as the basis vectors of the DCT-II matrix [44] such that  $\mathbf{U}^H = \mathbf{Q}$  is the DCT-III matrix, with entries  $Q_{m,n} = c(m)\sqrt{\frac{2}{N}} \cos\left(\frac{\pi m(n+0.5)}{N}\right)$ , for  $0 \leq m, n \leq N - 1$ , and constants  $c(0) = \frac{1}{\sqrt{2}}$  and  $c(m) = 1$  for  $m \geq 1$ , with corresponding distinct eigenvalues  $\lambda_m = 2 - 2 \cos\left(\frac{\pi m}{N}\right)$ ,  $m \in \{0, 1, \dots, N - 1\}$ . According to [45], a  $K$ -sparse signal sampled with the DCT matrix can be perfectly reconstructed via a variation of Prony's method using at least  $4K$  of its consecutive sample values which gives rise to a specialized extension of the Graph FRI framework<sup>4</sup>:

**Theorem 4.3.** (*Graph-FRI for paths*) Let  $\mathbf{x} \in \mathbb{C}^N$  be a  $K$ -sparse graph signal defined on the vertices of an undirected and unweighted path graph  $G$ , whose GFT basis is expressed such that  $\mathbf{U}^H$  is the DCT-matrix  $\mathbf{Q}$ . We can sample and perfectly reconstruct  $\mathbf{x}$  on  $G$  using the dimensionality-reduced GFT-representation  $\mathbf{y} = \mathbf{U}_M^H \mathbf{x} \in \mathbb{C}^M$ , where  $\mathbf{U}_M^H$  corresponds to the first  $M$  rows of  $\mathbf{U}^H$ , provided  $M \geq 4K$ .

Furthermore, we note that the graph Laplacian matrix of a path is circulant up to its first and last row, thus incorporating 2 vanishing moments with a graph-inherent border effect of length  $T = 2$ ; powers of the graph Laplacian similarly inherit  $2k$  vanishing moments with a border effect of  $T = 2k$ . By considering the symmetric normalized adjacency

<sup>3</sup>This can be easily demonstrated in a generalization of the proof of Thm. 3.2 (in Appendix A.2, [20]), where high-pass filter  $\mathbf{H}_{HP_{\vec{\alpha}}}$  is maintained and low-pass filter  $\mathbf{H}_{LP_{\vec{\alpha}}}$  is generalized to the form of eq. (6), with fixed downsampling pattern with respect to  $s = 1 \in S$ .

<sup>4</sup>In [45], the DCT matrix is given as an example of a larger class of invertible sampling bases of the form  $\mathbf{Q} = \mathbf{A}\mathbf{V}\mathbf{S}$ , with diagonal  $\mathbf{A} \in \mathbb{C}^{N \times N}$ , Vandermonde matrix  $\mathbf{V} \in \mathbb{C}^{N \times M}$  with  $[\mathbf{V}]_{n,m} = p_m^n$  and distinct  $p_m$ , and  $\mathbf{S} \in \mathbb{C}^{M \times N}$ , whose columns are at most  $D$ -sparse. A  $K$ -sparse signal  $\mathbf{x}$  can be perfectly recovered from  $2DK$  consecutive entries of  $\mathbf{y} = \mathbf{Q}\mathbf{x}$  using Prony's method (Prop. 4, [45]), which allows a further generalization to graphs whose GFT basis is of that form.

matrix  $\mathbf{A}_n = \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2}$ , the previously derived graph wavelet construction of Thm. 3.1 can be generalized for the path graph, even though not a regular graph, with the minor restriction that the normalized graph Laplacian  $\mathbf{L}_{norm} = \mathbf{I}_N - \mathbf{A}_n$  gains an increased border effect of  $T = 2(k + 1)$ . More generally, the proof of Thm. 3.1 can be extended to undirected connected graphs [20], given that the eigenvalues  $\gamma_i$  of  $\mathbf{A}_n$  continue to satisfy  $|\gamma_i| \leq 1$ . We thus state the graph wavelet transform as a special case of the HGSWT in Thm 3.1:

**Corollary 4.2.** *Given the undirected path graph  $G$  with normalized adjacency matrix  $\mathbf{A}_n$ , we define the HGSWT, composed of the low-and high-pass filters:*

$$\mathbf{H}_{LP} = \frac{1}{2^k} (\mathbf{I}_N + \mathbf{A}_n)^k \quad (12)$$

$$\mathbf{H}_{HP} = \frac{1}{2^k} (\mathbf{I}_N - \mathbf{A}_n)^k \quad (13)$$

*whose associated high-pass representer polynomial  $H_{HP}(z)$  annihilates polynomial graph signals up to degree  $2k - 1$ , subject to a graph-border effect of  $T = 2(k + 1)$  non-zeros. This filterbank is invertible for any downsampling pattern, as long as at least one node retains the low-pass component, while the complementary set of nodes retains the high-pass components.*

Hereby, the nullspace of  $\mathbf{L}_{norm}$  does not incorporate the all-constant vector, but rather the variation  $\mathbf{D}^{1/2} \mathbf{1}_N$ , with  $\mathbf{1}_N$  as the  $N$ -dimensional vector of all ones.

Similar extensions pertain to the e-graph-spline transform in Thm. 3.2, whose filters we can generalize to be of the form  $\mathbf{H} = \frac{1}{2^k} (\mathbf{I}_N \lambda_j \pm \mathbf{A}_n)^k$  with respect to the eigenvalues  $\lambda_j$  of  $\mathbf{A}_n$ , in the spirit of the eigenspace-shift discussed in Sect. 3, however, we are less interested in these results, except in order to achieve a sparse multiresolution representation, given that the DCT does not give rise to an equivalently intuitive decomposition scheme as the DFT for sampling-based graph coarsening. These graph filters are shift-invariant with respect to the normalized adjacency matrix since they are formed by polynomials in  $\mathbf{A}_n$ .

At last, it should be noted that generalized graph coarsening of a path graph, within a multilevel graph wavelet analysis, can i.a. be conducted via the Kron-reduction of the graph Laplacian matrix, whereby every other node is sampled, resulting in a weighted path graph with universal weight  $1/2$  [36], as well as via a spectral sampling scheme similar to Lemma 4.2, as a consequence of its near-circulant structure.

## 5. Generalized & Multidimensional Sparse Sampling

In order to apply the presented sampling framework to sparse signals defined on arbitrary graphs, one can make use of a variety of approximation schemes, which facilitate the interpretation of circulant graphs as building blocks for the former. Given the adjacency matrix  $\mathbf{A}$  of a general graph  $G$ , we propose to conduct (i) nearest circulant approximation of  $\mathbf{A}$  by  $\hat{\mathbf{A}}$ , or alternatively (ii) the (approximate) graph product decomposition  $\mathbf{A} \approx \mathbf{A}_1 \diamond \mathbf{A}_2$  into circulants  $\mathbf{A}_i$ .

In the former case, this entails the projection of  $\mathbf{A}$  onto the subspace of circulant matrices

$\mathbf{C}_N$ , spanned by circulant permutation matrices  $\mathbf{\Pi}^i$ ,  $i = 0, \dots, N-1$ , with  $\mathbf{\Pi}$  having first row  $[0 \ 1 \ 0 \dots]$ , via the Frobenius inner product  $\tilde{\mathbf{A}} = \sum_{i=0}^{N-1} \frac{1}{N} \langle \mathbf{A}, \mathbf{\Pi}^i \rangle_F \mathbf{\Pi}^i$  [46]. Hereby,  $\mathbf{A}$  may be subjected to a prior node relabelling for bandwidth minimization (using for instance the RCM-algorithm [47]) so as to reduce the approximation effect of introducing additional complementary edges in  $\tilde{\mathbf{A}}$  and hence significantly alter the graph, as well as prior partitioning, e.g. when the graph at hand features distinct communities. The set of (wavelet-)sparse signals defined on  $G$  can then be sampled with respect to its graph approximation  $\tilde{G}$ . In the latter case, we introduce an additional degree of dimensionality reduction through the graph product operation, which can be successfully leveraged for suitably defined sparse signals consisting of sparse tensor factors.

Analogously to our investigation of multidimensional graph wavelet analysis, outlined in [20], we can thus generalize the GFRI sampling framework to arbitrary graphs by resorting to graph product decomposition ([48], [49]); for completeness, we briefly review the main aspects.

The product  $\diamond$  of two graphs  $G_1 = (V(G_1), E(G_1))$  and  $G_2 = (V(G_2), E(G_2))$ , also referred to as factors, with respective adjacency matrices  $\mathbf{A}_1 \in \mathbb{R}^{N_1 \times N_1}$  and  $\mathbf{A}_2 \in \mathbb{R}^{N_2 \times N_2}$ , gives rise to a new graph  $G$  with vertex set  $V(G) = V(G_1) \times V(G_2)$  as the Cartesian product of the former, and edge set  $E(G)$  which is formed according to adjacency rules of the respective product operation, resulting in adjacency matrix  $\mathbf{A} \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$  [48]. We identify four main graph products of interest:

- the Kronecker product  $G_1 \otimes G_2$ :  $\mathbf{A}_{\otimes} = \mathbf{A}_1 \otimes \mathbf{A}_2$
- the Cartesian product  $G_1 \times G_2$ :  $\mathbf{A}_{\times} = \mathbf{A}_1 \times \mathbf{A}_2 = \mathbf{A}_1 \otimes \mathbf{I}_{N_2} + \mathbf{I}_{N_1} \otimes \mathbf{A}_2$
- the Strong product  $G_1 \boxtimes G_2$ :  $\mathbf{A}_{\boxtimes} = \mathbf{A}_1 \boxtimes \mathbf{A}_2 = \mathbf{A}_1 \otimes \mathbf{A}_2 + \mathbf{A}_1 \otimes \mathbf{I}_{N_2} + \mathbf{I}_{N_1} \otimes \mathbf{A}_2$
- the Lexicographic product  $G_1[G_2]$ :  $\mathbf{A}_{[\ ]} = \mathbf{A}_1[\mathbf{A}_2] = \mathbf{A}_1 \otimes \mathbf{J}_{N_2} + \mathbf{I}_{N_1} \otimes \mathbf{A}_2$

where  $\mathbf{J}_{N_2} = \mathbf{1}_{N_2} \mathbf{1}_{N_2}^T$ . Hereby, the lexicographic product is a special case of the Cartesian product. When  $G_i$  (and hence  $G$ ) are regular as well as connected, both adjacency  $\mathbf{A}_{\diamond}$  and graph Laplacian matrix  $\mathbf{L}_{\diamond}$  possess the same eigenbasis  $\mathbf{U} = \mathbf{U}_1 \otimes \mathbf{U}_2$  for  $\mathbf{A}_i = \mathbf{U}_i \mathbf{\Lambda}_i \mathbf{U}_i^{-1}$  on  $G_i$ , with graph adjacency eigenvalues  $\mathbf{\Lambda}_{\diamond} = \mathbf{\Lambda}_1 \otimes \mathbf{\Lambda}_2$ , except under the lexicographic product [50].

Inspired by the consideration of graph products in other contexts as a means to model higher-dimensional data and/or facilitate efficient implementation, and a preliminary consideration in [49] for GSP, we proceed to interpret a graph signal residing on the vertices of a product graph as follows [20]:

**Definition 5.1.** Any graph signal  $\mathbf{x} \in \mathbb{R}^N$ , with  $N = N_1 N_2$ , can be decomposed as  $\mathbf{x} = \sum_{s=1}^k \mathbf{x}_{s,1} \otimes \mathbf{x}_{s,2} = \text{vec}_r \{ \sum_{s=1}^k \mathbf{x}_{s,1} \mathbf{x}_{s,2}^T \}$ , whereby  $\text{vec}_r \{ \}$  indicates the row-stacking operation, or, equivalently,  $\sum_{s=1}^k \mathbf{x}_{s,1} \mathbf{x}_{s,2}^T$  has rank  $k$  with  $\mathbf{x}_{s,i} \in \mathbb{R}^{N_i}$ . For  $\mathbf{x}$  residing on the vertices of an arbitrary undirected graph  $G$ , which admits the graph product decomposition of type  $\diamond$ , such that  $G_{\diamond} = G_1 \diamond G_2$  and  $|V(G_i)| = N_i$ , we can redefine and process  $\mathbf{x}$  as the graph signal tensor factors  $\mathbf{x}_{s,i}$  on  $G_i$ .

Given the graph product decomposition  $G = G_1 \diamond G_2$ , which can be exact or approximate such that  $G_i$  are undirected, circulant and connected with  $s = 1 \in S_i, i = 1, 2$ , the tensor

factors of a given graph signal on  $G$  can thus be processed with respect to its inherent circulant substructures; for simplicity, we consider  $\mathbf{x} = \mathbf{x}_1 \otimes \mathbf{x}_2$  (at rank  $k = 1$ ) for the remainder of this discussion.

A  $K = K_1 K_2$ -sparse graph signal  $\mathbf{x}$ , with  $K_i$ -sparse tensors  $\mathbf{x}_i$ , on an arbitrary graph  $G$ , can be sampled and perfectly reconstructed based on dimensionality-reduced GFT-representations of  $\mathbf{x}_i$  on the approximate or exact graph product decomposition of  $G$  into circulant factors  $G_i$ . Hereby, one can apply the GFRI framework (Thms 4.1 & 4.2) on each component individually and thus perfectly recover  $\mathbf{x}_i$  (using Prony's method) from spectral representations  $\mathbf{y}_i = \mathbf{U}_{M_i}^H \mathbf{x}_i$ , of dimension  $M_i \geq 2K_i$ , with  $\mathbf{U}_{M_i}^H$  denoting the first  $M_i$  rows of the permuted GFT (DFT) matrix of dimension  $N_i \times N_i$ ,  $i = 1, 2$ . In particular, for all but the lexicographic product, we have  $(\Phi_{M_1} \otimes \Phi_{M_2}) \mathbf{U}^H \mathbf{x} = (\mathbf{U}_{M_1}^H \otimes \mathbf{U}_{M_2}^H)(\mathbf{x}_1 \otimes \mathbf{x}_2) = \mathbf{y}_1 \otimes \mathbf{y}_2 = \mathbf{y}$ , where  $\mathbf{y} \in \mathbb{C}^M$  with  $M \geq 4K$  and  $\Phi_{M_i} \in \mathbb{R}^{M_i \times N_i}$  sample the first  $M_i$  rows, or alternatively,  $\mathbf{y} = (\mathbf{C}_1 \tilde{\mathbf{y}}_1) \otimes (\mathbf{C}_2 \tilde{\mathbf{y}}_2)$  for graph-filtered representations  $\tilde{\mathbf{y}}_i$  and graph spectral transformation matrices  $\mathbf{C}_i$ . Otherwise, a separate processing of  $\mathbf{x}_i$  is conducted on the individual graph Laplacian eigenbases of  $G_i$ . The coarsened circulant graphs  $\tilde{G}_i$  associated with representations  $\tilde{\mathbf{y}}_i$  can be recombined under the same graph product operation to form the coarse graph  $\tilde{G} = \tilde{G}_1 \diamond \tilde{G}_2$  on the vertices of which the signal  $\tilde{\mathbf{y}}_1 \otimes \tilde{\mathbf{y}}_2$  is redefined.

### 5.1. Exact vs Approximate Graph Products

If the product decomposition  $G = G_1 \diamond G_2$  into circulants  $G_i$  is exact and known, we can perfectly recover a tensor-sparse signal  $\mathbf{x}$  on  $G$  by performing graph operations in smaller dimensions, while requiring the storage of lower dimensional spectral representations  $\mathbf{y}_i$ ; this advantage is particularly evident in the case of the lexicographic product which is closed under circulant graphs [51]. Hereby, one may apply the GFRI framework directly on the original  $G$ , with known lexicographic decomposition, conditional upon  $\mathbf{x}$  being sufficiently sparse (or compressible on  $G$ ), yet a decomposition into lower-dimensional circulants increases efficiency, while preserving the scheme.

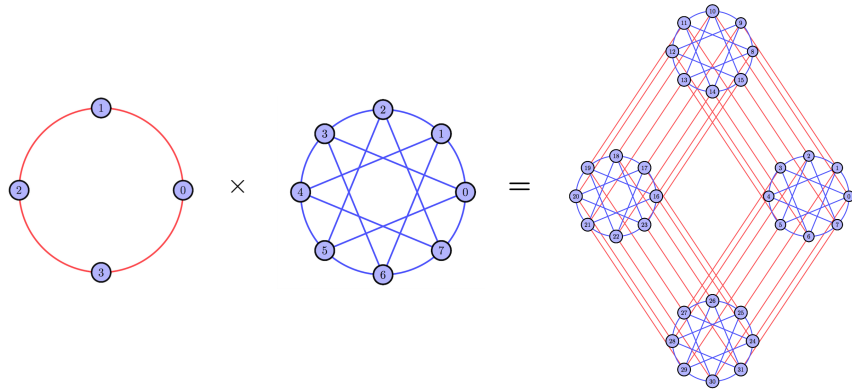


Figure 5: Graph Cartesian Product of two unweighted, bipartite circulant graphs.

The study of graph products has revealed that the Cartesian product of two unweighted

path graphs produces an unweighted lattice graph [48], facilitating the generalization of both the sparse sampling and wavelet analysis framework to lattice graphs, as well as to more general graph products of path and circulant graphs (see Fig. 5, 6 for examples). Due to the vanishing moment property of the graph Laplacian of a path graph, it becomes evident that one can apply a multidimensional wavelet analysis of compressible signals on lattice graphs; hereby, we note an interesting relation to the interpretation of the graph Laplacian as a differential operator. In particular, it is known that the graph Laplacian provides the stencil approximation of the second order differential operator for lattice graphs up to a sign [17]; coincidentally, unweighted lattice graphs, as the graph product of two path graphs, which are circulant up to a missing edge, imply that the inherent vanishing property of the graph Laplacian of a circulant graph is to some extent preserved via the product operation. We investigated this phenomenon further for general circulants in [20].

Following the GFRI Thm. for paths, one can equivalently perform multidimensional

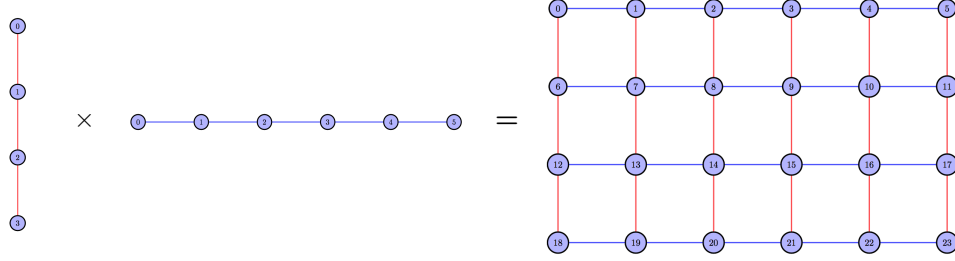


Figure 6: Graph Cartesian Product of two unweighted path graphs

sampling and reconstruction of signals defined on the graph product of path graphs, whereby sparse tensors  $\mathbf{x}_i$  on  $G_i$  can be perfectly recovered based on at least  $4K_i$  consecutive samples of their dimensionality-reduced GFT (DCT) representation.

#### 5.1.1. The Kronecker product approximation

In order to extend the GFRI framework beyond circulant and path graphs to arbitrary graphs, which can be decomposed as a graph product of the former, we resort to posing the decomposition in more general terms as an optimization problem subject to linear constraints and draw from a result in matrix theory [52]. The adjacency matrix  $\mathbf{A}$  of an arbitrary graph  $G$  can thus be approximated in Frobenius-norm as the Kronecker product  $\mathbf{A}_1 \otimes \mathbf{A}_2$  of (adjacency) matrices  $\mathbf{A}_i$  of dimension  $N_i$ , which are restricted to be circulant by imposing linear constraints:

$$\min_{\mathbf{C}_1^T \text{vec}(\mathbf{A}_1)=0, \mathbf{C}_2^T \text{vec}(\mathbf{A}_2)=0} \|\mathbf{A} - \mathbf{A}_1 \otimes \mathbf{A}_2\|_F.$$

Hereby,  $\mathbf{C}_i$  are structured, rectangular matrices with entries  $\{0, 1, -1\}$ , which can enforce symmetry and bandedness, in addition to circularity, via the column-stacking operation by  $\text{vec}$  [53]; the specific case of a 1-banded Toeplitz structure can similarly be imposed on either  $\mathbf{A}_i$  to obtain a decomposition into path graphs.

### 5.2. Multidimensional Separable Graph Wavelet Analysis

At last, we briefly revisit the *separable* graph wavelet transform, which is defined on the individual circulant factors  $G_i$  of product graphs, as a means to further extend the foregoing discussion on sampling. Let  $\mathbf{W}_i$  denote the graph (e-)spline wavelet transform constructed in the vertex domain of circulant graph factor  $G_i$ , as defined in Sect. 4.1, and  $\mathbf{w}_i = \mathbf{P}_{N_i} \mathbf{W}_i \mathbf{x}_i$  the graph wavelet domain representation of  $\mathbf{x}_i$  on  $G_i$ , subject to the (node relabelling) permutation  $\mathbf{P}_{N_i}$ . Then the multilevel transform

$$\mathbf{W}_i^{(j)} = \begin{bmatrix} \mathbf{W}_i^j & \\ & \mathbf{I}_{N_i - \frac{N_i}{2^j}} \end{bmatrix} \cdots \mathbf{W}_i^0$$

induces the representation  $\mathbf{w}_i = \mathbf{P}_{N_i}^{(j)} \mathbf{W}_i^{(j)} \mathbf{x}_i$ , with iterated permutation matrix

$$\mathbf{P}_{N_i}^{(j)} = \mathbf{P}_i^0 \cdots \begin{bmatrix} \mathbf{P}_i^j & \\ & \mathbf{I}_{N_i - \frac{N_i}{2^j}} \end{bmatrix}$$

at  $j \leq J - 1$  levels. The representation  $\mathbf{w} = \mathbf{w}_1 \otimes \mathbf{w}_2$  redefined on  $G$  is the result of a separable, two-dimensional graph spline wavelet transform, as introduced in [20].

If  $\mathbf{x}$  consists of compressible graph signal tensors  $\mathbf{x}_i$  such that their 2-D multiresolution graph wavelet representation  $\mathbf{w} = \mathbf{w}_1 \otimes \mathbf{w}_2 = \mathbf{P}_{N_1 N_2}^{(j)} (\mathbf{W}_1^{(j)} \otimes \mathbf{W}_2^{(j)}) \mathbf{x}$  is  $K$ -sparse with  $\|\mathbf{w}_i\|_0 = K_i$  and  $K = K_1 K_2$ , for suitable graph wavelet transforms  $\mathbf{W}_i^{(j)}$  at level  $j \leq J - 1$ , we can similarly apply the framework of sampling and perfect reconstruction of  $\mathbf{w}_i$  on the vertices of  $G_i$ . Eventually the original  $\mathbf{x}$  can be reconstructed from  $\mathbf{w}$ , subject to invertibility of the 2-D graph wavelet transform.

## 6. Conclusion

In this work, we have introduced a novel framework for the sampling and perfect reconstruction of sparse and wavelet-sparse graph signals with an associated graph coarsening scheme for circulant graphs, which can be generalised to arbitrary graphs, i.a. via graph product decomposition. Hereby, we have leveraged previously developed families of graph spline and graph e-spline wavelets which further facilitate the extension of our framework to the wider class of compressible graph signals. It would be of interest to explore further sparsifying transforms on graphs within our sampling framework by tackling the more generalized problem of identifying a suitable transform, given an arbitrary graph and graph signal, which can induce a sparse representation.

## Appendix A.

### Appendix A.1.

*Proof of Corollary 3.1.* Let  $\mathbf{W} = \frac{1}{2}(\mathbf{I}_N + \mathbf{K})\mathbf{H}_{LP} + \frac{1}{2}(\mathbf{I}_N - \mathbf{K})\mathbf{H}_{HP}$  represent the generalized *HGESWT*-matrix from Thm. 3.2 with  $\mathbf{H}_{LP/HP} = \prod_{n=1}^T \frac{1}{2^k} (\beta_n \mathbf{I}_N \pm \frac{\mathbf{A}}{d})^k$  and



diagonal downsampling matrix  $\mathbf{K}$ , with  $K(i, i) = 1$  at even-numbered (node) positions, and  $K(i, i) = -1$  otherwise. Then we obtain

$$\mathbf{W}^T \mathbf{W} = \frac{1}{2}(\mathbf{H}_{LP}^2 + \mathbf{H}_{HP}^2 + \mathbf{H}_{LP} \mathbf{K} \mathbf{H}_{LP} - \mathbf{H}_{HP} \mathbf{K} \mathbf{H}_{HP}) = \frac{1}{2}(\mathbf{H}_{LP}^2 + \mathbf{H}_{HP}^2)$$

whereby the RHS is the result of the equality  $\mathbf{K} \mathbf{H}_{LP} = \mathbf{H}_{HP} \mathbf{K}$  (and the equivalent in representer polynomial form  $H_{HP}(z) = H_{LP}(-z)$ ). Thus  $\mathbf{W}^T \mathbf{W}$  is circulant, i.e. it has the same basis as the circulant adjacency matrix  $\frac{\mathbf{A}}{d} = \mathbf{V} \mathbf{\Gamma} \mathbf{V}^H$  and its eigenvalues can be expressed as  $\Lambda = \frac{1}{2}(\prod_{n=1}^T \frac{1}{2^k} (\beta_n \mathbf{I}_N + \mathbf{\Gamma})^{2k} + \prod_{n=1}^T \frac{1}{2^k} (\beta_n \mathbf{I}_N - \mathbf{\Gamma})^{2k})$ . Hence, the condition number of  $\mathbf{W}$  is given by  $C = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$ .  $\square$

#### Appendix A.2.

*Proof of Corollary 4.1.* (i) The number of non-zero high-pass coefficients after applying one level of the HGSWT is  $B$ ; due to the additional ‘border effect’ of the low-pass filter at subsequent levels, we obtain the following series after  $j$  levels

$$S = B + \left(B + \frac{B}{2}\right) + \left(B + \frac{\frac{B}{2} + B}{2}\right) + \cdots = \sum_{n=0}^{j-1} (j-n) \frac{B}{2^n}.$$

Using the finite summation results

$$\sum_{n=0}^{j-1} \frac{j}{2^n} = j(2 - 2^{-j+1}), \quad \text{and} \quad \sum_{n=0}^{j-1} \frac{n}{2^n} = 2^{(1-j)}(-j - 1 + 2^j)$$

and considering the  $\frac{N}{2^j}$  low-pass coefficients, we obtain  $K = \frac{N}{2^j} + B(2(j-1) + 2^{1-j})$  as the total number of non-zeros. Hereby, for large  $B$ , the number of high-pass coefficients at each level  $l \leq j-1$  is bounded  $\sum_{n=0}^l \frac{B}{2^n} \leq \frac{N}{2^{l+1}}$ . If  $B = 2^{j-1}r$ , the formula for  $S$  is exact, otherwise, since  $S$  has to be an integer, we need to adjust the formula by adding/subtracting a term  $s_l$  at each level  $l$ , depending on whether downsampling requires rounding up or down. In particular, at each level, the high-pass filter is applied on the odd-numbered nodes  $1, 3, \dots$  of the (previously) low-pass filtered and sampled graph signal  $\tilde{\mathbf{y}}$ ; we thus note that if the length of the non-zero ‘border’ support (before downsampling) of the high-pass filtered  $\tilde{\mathbf{y}}$  at the beginning of the resulting labelled sequence is an even number, while that at the end of the sequence is odd, we need to round up, and vice versa.

(ii) Following the reasoning of the previous proof, we need to consider the border effect caused by filtering with the low-pass filter of bandwidth  $T$ ; we therefore end up with the following series summation for the total number of non-zeros

$$K = \frac{N}{2^j} + B + \left(B + \frac{T}{2}\right) + \left(B + \frac{\frac{T}{2} + T}{2}\right) + \cdots = \frac{N}{2^j} + jB + \sum_{n=1}^{j-1} \frac{T(j-n)}{2^n},$$

giving the formula  $K = \frac{N}{2^j} + jB + T(j-2 + 2^{1-j})$  subject to a correction term  $\pm s_l$  per level.

(iii) By Thm 3.1, we need to retain at least one low-pass component for invertibility of the filterbank, therefore we choose to assign the low-pass component to only one node, while the remaining nodes retain the high-pass components. While this downsampling approach is not conducted with respect to the generating set of the circulant graph, and therefore less rigorous from a graph-theoretical perspective, it achieves a maximally sparse representation in the graph wavelet domain. The number of non-zeros are  $2B$ , whereby  $2B - 1$  is the number of non-zero high-pass coefficients.  $\square$

### Appendix A.3.

*Proof of Thm. 4.1.* (Prony's method [45]):

Given the representation  $\mathbf{y} = \mathbf{U}_M^H \mathbf{x}$ , where  $\|\mathbf{x}\|_0 = K$  and  $\mathbf{U}_M^H$  are the first  $M$  rows of the DFT-matrix, we can represent the  $n$ -th entry of  $\mathbf{y}$  as  $y_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{K-1} x_{c_k} e^{-i2\pi c_k n/N}$  with weights  $x_{c_k}$  of  $\mathbf{x}$  at positions  $c_k$ , and apply Prony's method to recover  $\mathbf{x}$ , provided  $M \geq 2K$ . Hereby, we redefine  $y_n = \sum_{k=0}^{K-1} \alpha_k u_k^n$  with locations  $u_k = e^{-i2\pi c_k n/N}$  and amplitudes  $\alpha_k = x_{c_k}/\sqrt{N}$ , which are successively recovered. In the following, we summarise the reconstruction algorithm: given the samples  $y_n$ , we construct a Toeplitz matrix  $\mathbf{T}_{K,l}$ , and determine the vector  $\mathbf{h}$ , which lies in its nullspace, also known as 'the annihilating filter':

$$\mathbf{T}_{K,l} \mathbf{h} = \begin{pmatrix} y_{l+K} & y_{l+K-1} & \dots & y_l \\ y_{l+K+1} & y_{l+K} & \dots & y_{l+1} \\ \vdots & \ddots & \ddots & \vdots \\ y_{l+2K-2} & \ddots & \ddots & \vdots \\ y_{l+2K-1} & y_{l+2K-2} & \dots & y_{l+K-1} \end{pmatrix} \begin{pmatrix} 1 \\ h_1 \\ h_2 \\ \vdots \\ h_K \end{pmatrix} = \mathbf{0}$$

which can be accomplished via the SVD-decomposition of  $\mathbf{T}_{K,l}$ . Hereby, it can be shown that  $\mathbf{T}_{K,l}$  is of rank  $K$  for distinct  $u_k$  (Prop. 1, [45]). In particular, this corresponds to the matrix-form expression of  $\sum_{0 \leq k \leq K-1} \alpha_k u_k^n P(u_k) = 0$  for  $l \leq n < l + K$ , with polynomial

$$P(x) = x^K + \sum_{n=1}^K h_n x^{K-n} = \prod_{k=1}^K (x - u_{k-1})$$

whose roots  $\{u_k\}_{k=0}^{K-1}$  can be subsequently determined from  $\mathbf{h}$ . At last, we can recover the corresponding amplitudes  $\{\alpha_k\}_{k=0}^{K-1}$  by solving a system of  $K$  linear equations given by  $y_n$ .  $\square$

### Appendix A.4.

*Proof of Lemma 4.2.* (appears in part in [19]).

The eigenvalues of  $\mathbf{A}$  with first row  $[0 \ a_1 \dots a_1]$  are  $\lambda_j = \sum_{k=1}^B 2a_k \cos\left(\frac{2\pi j k}{N}\right)$ ,  $j = 0, \dots, N-1$ . Thus the eigenvalues of  $\tilde{\mathbf{A}}$  with the same entries  $a_i$  and bandwidth  $B < N/4$ , are  $\tilde{\lambda}_j = \sum_{k=1}^B 2a_k \cos\left(\frac{2\pi(2j)k}{N}\right) = \lambda_{2j}$ ,  $j = 0, \dots, N/2-1$ . We can similarly show the preservation of the downsampled DFT-eigenbasis. Let

$$\mathbf{x} = [e^{i\alpha 0} \ e^{i\alpha 1} \ e^{i\alpha 2} \ \dots \ e^{i\alpha(N-1)}]^T$$

with  $\alpha = -\frac{2\pi k}{N}$  denote the  $(k+1)$ -th row of the non-normalized DFT-matrix. If we discard all entries at odd-numbered positions, we obtain the  $(k+1)$ -th row of the DFT of dimension  $N/2$ , since

$$[e^{i\alpha 0} \ e^{i\alpha 2} \ e^{i\alpha 4} \ \dots \ e^{i\alpha(N-2)}]^T = [e^{i(2\alpha)0} \ e^{i(2\alpha)1} \ e^{i(2\alpha)2} \ \dots \ e^{i(2\alpha)(N/2-1)}]^T$$

with  $2\alpha = -\frac{2\pi k}{N/2}$ . Thus, if we apply the above sampling pattern on the first  $N/2$  rows of the DFT of dimension  $N$ , we obtain the DFT of dimension  $N/2$ . In particular, at  $k = N/2$ , we have  $\alpha = \pi$  and thus corresponding, downsampled row  $\mathbf{x} = \mathbf{1}_{N/2}$ , and proceeding similarly, we observe that the sampled lower half of the  $N \times N$  DFT, equivalently gives the DFT of dimension  $N/2$ . □

## References

- [1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, P. Vandergheynst, The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains, *IEEE Signal Process. Mag.* 30 (3) (2013) 83–98.
- [2] F. R. K. Chung, *Spectral Graph Theory*, American Mathematical Society, 1997.
- [3] M. Püschel, J. M. F. Moura, Algebraic signal processing theory: Foundation and 1-d time, *IEEE Transactions on Signal Processing* 56 (8) (2008) 3572–3585. doi:10.1109/TSP.2008.925261.
- [4] G. H. Golub, C. F. V. Loan, *Matrix computations* (3rd ed.), Johns Hopkins University Press, Baltimore, MD, USA, 1996.  
URL <http://portal.acm.org/citation.cfm?id=248979>
- [5] R. Coifman, M. Maggioni, Diffusion wavelets, *Applied and Computational Harmonic Analysis* 21 (1) (2006) 53–94.
- [6] V. N. Ekambaram, G. Fanti, B. Ayazifar, K. Ramchandran, Critically-sampled perfect-reconstruction spline-wavelet filterbanks for graph signals, in: *IEEE Global Conference on Signal and Information Processing, GlobalSIP 2013*, 2013, pp. 475–478. doi:10.1109/GlobalSIP.2013.6736918.
- [7] S. K. Narang, A. Ortega, Perfect reconstruction two-channel wavelet filter banks for graph structured data, *Signal Processing, IEEE Transactions on* 60 (6) (2012) 2786–2799. doi:10.1109/TSP.2012.2188718.
- [8] S. K. Narang, A. Ortega, Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs, *Signal Processing, IEEE Transactions on* 61 (19) (2013) 4673–4685. doi:10.1109/TSP.2013.2273197.
- [9] D. K. Hammond, P. Vandergheynst, R. Gribonval, Wavelets on graphs via spectral graph theory, *Applied and Computational Harmonic Analysis* 30 (2) (2011) 129–150. doi:http://dx.doi.org/10.1016/j.acha.2010.04.005.  
URL <http://www.sciencedirect.com/science/article/pii/S1063520310000552>
- [10] S. K. Narang, A. Gadde, A. Ortega, Signal processing techniques for interpolation in graph structured data, in: *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2013*, Vancouver, BC, Canada, May 26–31, 2013, 2013, pp. 5445–5449. doi:10.1109/ICASSP.2013.6638704.  
URL <http://dx.doi.org/10.1109/ICASSP.2013.6638704>
- [11] X. Wang, P. Liu, Y. Gu, Local-set-based graph signal reconstruction, *Signal Processing, IEEE Transactions on* 63 (9) (2015) 2432–2444.
- [12] S. Chen, A. Sandryhaila, J. Moura, J. Kovačević, Signal recovery on graphs: Variation minimization, *IEEE Transactions on Signal Processing*.
- [13] S. K. Narang, Y. H. Chao, A. Ortega, Graph-wavelet filterbanks for edge-aware image processing, in: *IEEE Statistical Signal Processing Workshop, SSP 2012*, Ann Arbor, MI, USA, August 5–8, 2012, 2012, pp. 141–144. doi:10.1109/SSP.2012.6319643.  
URL <http://dx.doi.org/10.1109/SSP.2012.6319643>
- [14] A. Gadde, A. Anis, A. Ortega, Active semi-supervised learning using sampling theory for graph signals, in: *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2014, pp. 492–501.

- [15] V. N. Ekambaram, G. Fanti, B. Ayazifar, K. Ramchandran, Wavelet-regularized graph semi-supervised learning, in: Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE, IEEE, 2013, pp. 423–426.
- [16] V. N. Ekambaram, G. C. Fanti, B. Ayazifar, K. Ramchandran, Circulant structures and graph signal processing, in: IEEE International Conference on Image Processing, ICIP 2013, 2013, pp. 834–838. doi:10.1109/ICIP.2013.6738172.
- [17] V. Ekambaram, Graph structured data viewed through a fourier lens, Ph.D. thesis, EECS Department, University of California, Berkeley (Dec 2013).
- [18] M. S. Kotzagiannidis, P. L. Dragotti, Higher-order graph wavelets and sparsity on circulant graphs, in: SPIE Optical Engineering+ Applications, Vol. 9597, International Society for Optics and Photonics, 2015, pp. 95971E–95971E–9. doi:10.1117/12.2192003.
- [19] M. S. Kotzagiannidis, P. L. Dragotti, The graph fri framework-spline wavelet theory and sampling on circulant graphs, in: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 6375–6379. doi:10.1109/ICASSP.2016.7472904.
- [20] M. S. Kotzagiannidis, P. L. Dragotti, Splines and wavelets on circulant graphs (submitted), <http://arxiv.org/abs/1603.04917>.
- [21] M. Vetterli, P. Marziliano, T. Blu, Sampling signals with finite rate of innovation, IEEE Transactions on Signal Processing 50 (6) (2002) 1417–1428. doi:10.1109/TSP.2002.1003065.
- [22] P. L. Dragotti, M. Vetterli, T. Blu, Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets strang-fix, IEEE Trans. Signal Process. 55 (5) (2007) 1741–1757. doi:10.1109/TSP.2006.890907.
- [23] D. Thanou, D. I. Shuman, P. Frossard, Parametric dictionary learning for graph signals, in: IEEE Global Conference on Signal and Information Processing, GlobalSIP 2013, Austin, TX, USA, December 3-5, 2013, 2013, pp. 487–490. doi:10.1109/GlobalSIP.2013.6736921. URL <http://dx.doi.org/10.1109/GlobalSIP.2013.6736921>
- [24] A. G. Marques, S. Segarra, G. Leus, A. Ribeiro, Sampling of graph signals with successive local aggregations, IEEE Transactions on Signal Processing 64 (7) (2016) 1832–1843. doi:10.1109/TSP.2015.2507546.
- [25] I. Z. Pesenson, Sampling in paley-wiener spaces on combinatorial graphs, Trans. Amer. Math. Soc. 360 (10) (2008) 5603–5627.
- [26] A. Anis, A. Gadde, Akshay, A. Ortega, Towards a sampling theorem for signals on arbitrary graphs, in: Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, IEEE, 2014, pp. 3864–3868.
- [27] S. Chen, R. Varma, A. Sandryhaila, J. Kovačević, Discrete signal processing on graphs: Sampling theory, IEEE Transactions on Signal Processing 63 (24) (2015) 6510–6523. doi:10.1109/TSP.2015.2469645.
- [28] P. Liu, X. Wang, Y. Gu, Coarsening graph signal with spectral invariance, in: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2014, Florence, Italy, May 4-9, 2014, 2014, pp. 1070–1074. doi:10.1109/ICASSP.2014.6853761. URL <http://dx.doi.org/10.1109/ICASSP.2014.6853761>
- [29] F. Dorfler, F. Bullo, Kron reduction of graphs with applications to electrical networks, Circuits and Systems I: Regular Papers, IEEE Transactions on 60 (1) (2013) 150–163.
- [30] M. S. Kotzagiannidis, P. L. Dragotti, Sparse graph signal reconstruction and image processing on circulant graphs, in: IEEE GlobalSIP, 2014, pp. 923–927. doi:10.1109/GlobalSIP.2014.7032255.
- [31] D. L. Donoho, Compressed sensing, Information Theory, IEEE Transactions on 52 (4) (2006) 1289–1306.
- [32] S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (1) (2001) 129–159. doi:10.1137/S003614450037906X. URL <http://dx.doi.org/10.1137/S003614450037906X>
- [33] E. J. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Transactions on Information Theory 52 (2) (2006) 489–509. doi:10.1109/TIT.2005.862083.
- [34] P. Stoica, R. L. Moses, Introduction to spectral analysis, Vol. 1, Prentice hall Upper Saddle River, 1997.
- [35] E. J. Candes, Y. C. Eldar, D. Needell, P. Randall, Compressed sensing with coherent and redundant dictionaries, Applied and Computational Harmonic Analysis 31 (1) (2011) 59–73.
- [36] D. I. Shuman, M. J. Faraji, P. Vandergheynst, A multiscale pyramid transform for graph signals, arXiv preprint arXiv:1308.4942.
- [37] C. Vonesch, T. Blu, M. Unser, Generalized daubechies wavelet families, Signal Processing, IEEE

- Transactions on 55 (9) (2007) 4415–4429. doi:10.1109/TSP.2007.896255.
- [38] M. Vetterli, J. Kovačević, V. K. Goyal, Foundations of Signal Processing, Cambridge University Press, <http://www.fourierandwavelets.org>, 2014.
  - [39] J. A. Uriguen, T. Blu, P. L. Dragotti, Fri sampling with arbitrary kernels, IEEE Transactions on Signal Processing 61 (21) (2013) 5310–5323. doi:10.1109/TSP.2013.2278152.
  - [40] T. Blu, M. Unser, Approximation error for quasi-interpolators and (multi-)wavelet expansions, Applied and Computational Harmonic Analysis 6 (2) (1999) 219 – 251. doi:<http://dx.doi.org/10.1006/acha.1998.0249>  
URL <http://www.sciencedirect.com/science/article/pii/S1063520398902494>
  - [41] G. F. G. Strang, A fourier analysis of the finite element variational method, in: Constructive Aspects of Functional Analysis, Rome, Italy: Edizioni Cremonese, 1973, pp. 795–840.
  - [42] M. Unser, Ten good reasons for using spline wavelets, in: Proc. SPIE vol. 3169, Wavelet Applications in Signal and Image Processing V, 1997, pp. 422–431.
  - [43] M. Unser, T. Blu, Cardinal exponential splines: Part i—theory and filtering algorithms, IEEE Trans. Signal Process 53 (2005) 1425–1438.
  - [44] G. Strang, The discrete cosine transform, SIAM Review 41 (1999) 135–147.
  - [45] P. L. Dragotti, Y. M. Lu, On sparse representation in fourier and local bases, IEEE Transactions on Information Theory 60 (2014) 7888–7899. doi:10.1109/TIT.2014.2361858  
URL <http://dx.doi.org/10.1109/TIT.2014.2361858>
  - [46] T. F. Chan, An optimal circulant preconditioner for toeplitz systems, SIAM journal on scientific and statistical computing 9 (4) (1988) 766–771.
  - [47] E. Cuthill, J. McKee, Reducing the bandwidth of sparse symmetric matrices, in: Proceedings of the 1969 24th National Conference, ACM '69, ACM, 1969, pp. 157–172. doi:10.1145/800195.805928.  
URL <http://doi.acm.org/10.1145/800195.805928>
  - [48] R. Hammack, W. Imrich, S. Klavzar, Handbook of Product Graphs, Second Edition, 2nd Edition, CRC Press, Inc., Boca Raton, FL, USA, 2011.
  - [49] A. Sandryhaila, J. M. F. Moura, Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure, IEEE Signal Process. Mag. 31 (5) (2014) 80–90.
  - [50] H. Sayama, Estimation of laplacian spectra of direct and strong product graphs, CoRR abs/1507.03030.
  - [51] R. Hoshino, Independence polynomials of circulant graphs, Ph.D. thesis, Dalhousie University (2008).
  - [52] C. V. Loan, N. Pitsianis, Approximation with kronecker products, in: Linear Algebra for Large Scale and Real Time Applications, Kluwer Publications, 1993, pp. 293–314.
  - [53] N. P. Pitsianis, The kronecker product in approximation and fast transform generation, Ph.D. thesis, Ithaca, NY, USA, uMI Order No. GAX97-16143 (1997).